From national accounting to the design, compilation, and use of bayesian policy and analysis frameworks

Van Tongeren, J.W.

Publication date:
2011

Jan W. van Tongeren

From national accounting to the design, compilation, and use of Bayesian policy analysis frameworks

From national accounting to the design, compilation, and use of Bayesian policy analysis frameworks

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 14 oktober 2011 om 10.15 uur door

Jan Willem van Tongeren,

geboren op 12 september 1940 te Deventer.

PROMOTOR: PROF. DR. J.R. MAGNUS

To my wife Julia
To my children Kai en Ica
To the memory of my parents

Preface

This thesis is based on the following five papers:

Chapter 2: Van Tongeren, J.W. (2004), Designing analytical data frameworks, *The Review of Income and Wealth*, 50, 279–297.

Chapter 3: Bartelmus, P., C. Stahmer, and J.W. van Tongeren (1991), Integrated environmental and economic accounting: Framework for a SNA satellite system, *The Review of Income and Wealth*, 37, 111–148.

Chapter 4: Van Tongeren, J.W. (1986), Development of an algorithm for the compilation of national accounts and related systems of statistics, *The Review of Income and Wealth*, 32, 25–47.

Chapter 5: Magnus, J.R., J.W. van Tongeren, and A.F. de Vos (2000), National accounts estimation using indicator ratios, *The Review of Income and Wealth*, 46, 329–350.

Chapter 6: Van Tongeren, J.W. and J.R. Magnus (2011), Bayesian integration of large SNA data frameworks with an application to Guatemala, CentER Discussion Paper 2011-022.

Acknowledgements

The opportunities provided by the institutions in which I was employed and the ideas of colleagues and friends to which I was exposed during a 40-year period were important factors that influenced the direction of the present research.

I am most grateful for the opportunities provided by successive employers to implement intermediate and final results of the research. My first employment at the Stichting Planbureau Suriname provided me with the opportunity to become a leading member of an Inter-Departmental Working Group developing national accounts for Surinam, and thus enter the area of national accounting as amateur. My knowledge of national accounting concepts and the methods that were used to compile those was deepened and made more professional during the first years of my employment with the United Nations Statistics Division (UNSD). Thereafter, during my frequent travels, I was able to understand more of the intricacies of the compilation methods and gradually was able to make my own contributions to those. The 10-year process of the 1993 revision of the international guidelines on national accounting (1993 SNA) gave me insights into how concepts were developed and how they relate to each other within frameworks and how the national accounts methodology could be used in alternative frameworks. The years thereafter I was given the opportunity as interregional adviser to make my contributions to the 1993 SNA implementation in different countries. Also during those periods of employment I was given the opportunity to confront this knowledge with others in international organizations and countries, who were more interested in the policy uses of national accounts data. Furthermore, through a long-standing relationship with the Development Research Institute (IVO) at Tilburg University I was provided with an academic forum for discussions and presentations of SNA related issues. This also included the implementation of a Bayesian estimation approach to SNA compilation in several (mainly developing) country projects, and in one case (Rwanda) the design and implementation of a university curriculum for training in SNA and other frameworks and Bayesian compilation thereof.

Of course, these opportunities are not abstract conditions, but they are reflected in persons, with whom I have interacted professionally as colleagues and in many instances as friends. Their number is very large, and only some I will be able to thank explicitly through these lines, but many others have equally contributed to my ideas, and to those I am equally indebted.

During my Surinam years, I am grateful to Ronald Schermel, as Director General of the Ministry of Economic Affairs, and as colleague and friend for his support of the national accounting work in Surinam, to Hugo Rijsdijk and Jules de Goede, respectively official of the Rekenkamer (Government Auditing Office) and director of the Statistical Office of Surinam, for their technical support, to Humphrey Lamur as colleague at the Stichting Planbureau Suriname (Project Planning Institution) for in-depth knowledge of the Surinam society, and which he shared with me as a friend, and to Jaap Guyt, also as colleague and friend, for sharing with me his experiences in international cooperation.

I am very grateful to my former colleagues at UNSTAT (former name of UNSD) for the stimulating discussions that we had about national accounting concepts and their implementation. I should refer here for the beginning period of my UN employment, with much appreciation to Abraham Aidenoff, my boss at that time, who was with Richard Stone the main author of the 1968 SNA, to my Norwegian colleagues Hans Pedersen and Rainar Oines, both with much international experience and who had been involved in the effort by Richard Stone to develop national accounting after the Second World War, to my Russian colleague Yuri Ivanov, who still is a prominent expert on national accounting in the former Centrally Planned Economies, and to Barbara Campbell from Australia who had supported the development of the 1968 SNA. I also should mention here my appreciation to the late Kees Oomens, Director of the Netherlands CBS at that time and also member of the 1968 SNA Expert Group, who arranged for me a period of practice and training at the Netherlands CBS (Statistics Netherlands), as interruption of my UN employment, during which time I was to learn more about one of the most advanced systems of national accounts in the world at that time.

I remember with much appreciation my time in Bolivia, during which I got the first opportunity to implement national accounting in a country setting. In this I got much support from William Abraham, the team leader, Jaime Robles as the expert in Industrial Statistics, Priscila de Loayza who was the public sector statistics expert in the Instituto Nacional de Estadística (INE) in La Paz, and Freddy Justiniano, who was the project counterpart and later became Director of INE, and became Minister of Economic Affairs after I had left Bolivia.

After my return from Bolivia to UNSD at New York, I remember with gratitude the many stimulating discussions with Nancy and Richard Ruggles, among others in their home at New Haven, about the features of a future SNA in preparation for the next revision of the international guidelines, and also about the micro-macro link that should support future compilations of national accounts. I remember with envy the encyclopedic memory of Ruben Baratz who initiated the work on classifications at UNSD, and remember with appreciation the subsequent discussions on classifications with my UNSD colleagues Michael Beekman, Jim Barrat, Amy Ferrara, Mary Chamie and Ralph Becker, which have very much influenced my thoughts on classifications and correspondences. Equally interesting were my discussions with Lazlo Drechsler on the national accounting base of International Price Comparisons.

I remember with appreciation the many sessions that we had around the same time, with Derek Blades of OECD and Brian Newson of EUROSTAT, in which SNA revision work was organized. I should also mention in this context the support I received from my colleague and friend Cristina Hannig, with whom I shared many of my administrative functions as head of the National Accounts Branch, and which allowed me to dedicate more time to the technical aspects of national accounting. The SNA revision process that followed thereafter for a period of nearly 10 years was a rich ground on which I could build my knowledge and experiences of SNA and satellite concepts within a framework setting. The stimulating discussions in expert groups and ISWGNA (Inter-Secretariat Working Group on National Accounts) with colleagues of other organizations contributed much to my in-depth knowledge of national accounting concepts and frameworks: I refer here in particular with appreciation to the prominent contributions therein made by outstanding experts in this field, including André Vanoli, Peter Hill, Derek Blades, Brian Newson, Anne Harrison, Carol Carson, Kevin O'Connor, Pablo Mandler, Datuk Chander, Heinrich Lutzel, Jagdish Kumar, Uma Dat Choudhury, Mofat Nyoni, and Steven Keuning, who prepared the SNA chapter on Social Accounting Matrices (SAM's).

In the development of an operational approach to satellite accounting, I remember the many stimulating inter-disciplinary discussions on environmental accounting with Peter Bartelmus and Alessandra Alfieri of UNSD, Carsten Stahmer of the German Statistisches Bundesambt, Robert Repetto of the World Resources Institute, Ernst Lutz of the World Bank and many more, and how we managed as a team to implement environmental accounts in three selected countries. In later efforts to implement environmental accounting, I have benefited from the support by Guillermo Rudas in Colombia, Renato Vargas and Pedro Pineda in Guatemala, and my colleague Roldan Muradian at IVO. In the same vein, I would like to mention my appreciation about the interesting discussions I had with Helen Stone Tice and Lester Salamon of the John Hopkins University Center for Civil Society Studies about the application of national accounting standards to the sector of Non-Profits, to the Health sector with Ruben Suarez of the Pan-American Health Organization, and with my colleagues Bernd Becker and Guadalupe Espinosa, when preparing the materials for the Handbook on the household sector and other social dimensions of national accounting.

To my colleague Vu Viet I am much indebted for his contributions to my knowledge on Input-Output, applications of SNA to Business Accounting and many other types of analyses that could be based on SNA statistical standards. I am also grateful for my cooperation with Douglas Walker, an econometrician in the Policy Analysis Division of the UN, with whom I discussed many forms of analysis of the accounts and prepared joint papers. I am also very appreciative of the members of an Expert Group on the Use of SNA in Policy Analysis, which included among others Lawrence Klein, Graham Pyatt, Michael Ward, Thijs ten Raa, Carlos Blanco, Michael Ward and John Dawson.

In the development of the Bayesian approach to national accounts compilation, I am very much indebted to Jan Magnus of Tilburg University for his intellectual guidance and technical contributions to the present research, and for his development of the SNAER software programme, together with Dmitry Danilov and Judith Lammers, operationalizing the Bayesian estimation approach. I also would like to thank with much appreciation Arthur Giesberts, who not only contributed the Interface software programme, but also was the key player in the development of projects supporting the application of the Bayesian estimation approach to SNA implementation in various countries, and provided his unconditional support of this research, even after he left IVO and took up a different post at the Netherlands CBS.

Implementation of the compilation and framework methodologies developed is extremely important. This was done through the development of the System's approach to national accounting. This approach was developed after many discussions with experts on national accounts compilation, including Evaristo Arrieta, and Bernardo and Lourdes Ferrán. I am particularly grateful in this to Magda Ascues and Stefan Schweinfest. Magda implemented this approach in many countries in Central and Latin America and Africa, refined the methods used therein, and more recently shared her experiences on compilation in the further development of the Bayesian estimation approach. Stefan provided the support of the Systems Approach at UNSD in New York and also participated in the application of this approach to environmental accounting. In this context I also benefited much from the discussions with Michel Seruzier, who had much influence on national accounts compilation through the development of the ERETES software. Furthermore, I am very grateful to all those colleagues in a large number of countries that I visited, who had so much confidence in my efforts to apply this methodology in their institutions. Without being exhaustive of the many with whom I had contact and who contributed to the development of my ideas in national accounting related fields, I may mention a few. I benefited much from my cooperation with Carlos Blanco of the Consejo Monetario de America Central, Olga Diaz and Ian Abud from the Dominican Republic, Victor Flores and Rocío Lopez from Guatemala, Henry Vargas from Costa Rica and others who supported the implementation in a

project in Central America, which actually led to an operational instrument of framework design and Bayesian compilation. Bertha Vallejo, as project manager of the Central American project contributed much to the success of the Bayesian estimation in Central America, and at the same time contributed to university training in this approach in a project in Rwanda. I also benefited much from an experimental implementation of the Bayesian approach in the Netherlands CBS, under supervision of Brugt Kazemier, with whom I had many interesting discussions. Others provided the support in Latin America and the Caribbean to the System's and/or Bayesian approach to SNA implementation, including Carmen Reyes (Peru), Maria Eugenia Gómez Luna (Mexico), Vera Pérez (Colombia), Maureen Blokland (Netherlands Antilles), Hazel Corbin (Organization of East Caribbean States), Eneas Avondoglio (Argentina), José Venegas (Chile), Raúl García Belgrano (Peru and Economic Commission for Latin America and the Caribbean). In Africa I am grateful for the support from René Rakotobe (Economic Commission for Africa), Saíd Dade (Mozambique), Flavio Couto (Angola), and Johan Prinsloo and Herman Smit (South Africa). My thanks also go to Kotb Salem (Economic and Social Commission for Western Asia) and Salima Al Harthy (Oman), for supporting the implementation in the Arab countries, to Antonio Lazo, who supported the approach in Mozambique, Sao Tome en Principe, Oman, and lately in Peru, and to Jan Bartlema, who supported the approach in the Kurdistan region of Iraq. Also deserve many thanks, Heidi Arboleda, Estrella Domingo, Jagdish Kumar and Tiwari and Andy Flatt, who supported the approach in Asia.

A special thanks goes to Ruud Picavet, with whom I cooperated during a large part of my career, from the time he was still employed at Tilburg University and IVO and I was still at UNSD, until the present. What I appreciated most was his belief in the need to bring national accounting theory and practices into the academic sphere of a university. With this in mind I gave incidental lectures to his students at Tilburg University and he tried to interest students in the topics of this research. Special mention should be made of his contributions to the development of training materials supporting in particular a university curriculum in a Rwanda project. Furthermore, his continued readiness to critically review papers and maintain discussions about the Bayesian compilation approach has very much contributed to its operational implementation. His unconditional support, even in times that I nearly gave up, has very much contributed to where I am now in this field of research.

Very special thanks goes to Julia Santander, who as colleague, friend and presently wife followed and supported the larger part of my thesis work. She gave not only moral support, but at several points in time she also provided technical support to the research, particularly based on her extensive IT knowledge of computer and softwares to handle large databases. Once we lived together, she inherited not only the good moments and the joy of living and working together, but also the many frustrations that accompanied my research efforts. However, Julia, without hesitation you continued your support, and the fulfillment of my ambitions at present could not have been realized without your firm belief that I would reach the final goal of my research. Many thanks!

Contents

Chapter 1

Introduction

1 Context of the thesis

The subsequent chapters of this thesis are a selection of the materials that represent a comprehensive research effort, which I have pursued throughout my nearly 50-year involvement in national accounting. The purpose of this section is to describe the links between the various research efforts reflected in Chapters 2–6 and the context in which this research took place over time. The research efforts were mainly in response to two broad questions, i.e.

1. How could the scope and detail of national accounts be adapted to serve better policy and analytical uses; and

2. How could the methods of compiling national accounts be formalized, so that they can be more easily explained to users of the estimates and more effectively related to econometric methods used in analysis.

Both questions were part of the discussions that took place when preparing for the 1968 and 1993 revisions of the SNA (System of National Accounts); see UNSD (1968) and United Nations et al. (1993). The first question became relevant at the time that revised SNA frameworks needed to be defined. The second question was raised, when computerization entered the field of national accounting and the large data bases underlying national accounts estimates were being processed increasingly through computerized methods. Both questions were posed independently of each other. However, they have become more and more intertwined since then, as the implementation of in particular the 1993 SNA framework required increasingly large data sets that needed to be reconciled and could only be managed well through computerized processing. Furthermore, the increasing demand for satellite frameworks in response to specialized analytical needs focused more and more attention on how to design such frameworks, and on the processing of very diverse data sets.

The present research started with a focus on national accounting, where the methods originated. It has ended up, however, in the development of a joint operational instrument of framework design and compilation approach, that can be used in policy analysis research in general, covering a much wider set of data and estimates than national accounting (economic, social, environmental, demographic and even technical data), as long as a multitude of meaningful relations can be defined within the framework of data and estimates. Below will be explained how the chapters with selected materials played a role in the sequence of research efforts and what other issues were raised at the time and are not dealt with in those chapters.

Bringing accounting closer to the needs of policy analysis for statistical information is very much the objective of the present research effort. It is reflected in the UNSD Handbook on Macro Accounting for Policy Analysis (UNSD, 2002a) of which the present author was the editor (see also Bos, 2009a, Part II). It is also underlying the design of the Bayesian SNA framework in Chapter 6, which emphasized in the SNA framework for use by Central Banks in Central America, the role of the IEA (Integrated Economic Accounts) as a means of bringing closer real sector data with financial data in support of monetary-fiscal policies. And it is also the motivation behind the design in Chapter 3 of a framework for environmental-economic analysis in response to growing concerns that SNA did not respond well to environmental issues. Drastic changes in the international standard SNA framework are not always necessary, however, as Van Tongeren (2000) shows that many analytical indicators of production (e.g. I-O coefficients, output per worker) and imports (e.g. terms of trade effect), of household behaviour (e.g. household propensity to consume), of an institutional nature (e.g. tax rates for sectors) are already supported by the standard international SNA framework. The above variety of policy uses can all be supported statistically with help of the joint instrument of framework design and compilation approach, which can be applied to virtually any type of data set. This orientation of the research effort is very much in line with the work of Stone in his support of the 1968 SNA. In his memorial lecture Stone (1984) clearly showed interest in linking accounting and modeling,

> ‘by organising our data in the form of accounts we can obtain a coherent picture of the stocks and flows, incomings and outgoings of whatever variables we are interested in . . . ’

and

> 'by combining facts and theories we can construct a model which when translated into quantitative terms will give us an idea of how the system under investigation actually works ... '

In an earlier study Stone (1955) developed several simple models, using the elements of the accounts as a means of deriving relations between variables of the accounts, which then could be used in the estimation of the accounts for a coming year and also to assess the impact of changes in the economy.

The practices of national accounting and modeling, however, have gone in different directions. National accounting now focuses on the past and modeling is used to make projections to the future. Furthermore, modeling generally requires the estimation of parameters in multiple regression functions with help of time series, while national accounts uses structural binary coefficients of a base year. Also, national accounting is based on an explicit framework with ratios and identities holding between the estimates, while modeling does not adhere to an explicit framework, and therefore projections to the future of models do not necessarily satisfy all national accounts identities, partly because not all variables of those national accounts identities are included in the model. There may also be a conflict between the use of ratios in compilation and analysis, when conventional national accounting fixes the value of ratios on the basis of benchmark year data, which then become unusable for analysis. Thus I-O ratios between output and value added are often used in national accounting to make estimates of variables for which no basic data are available, but then their use in analysis is not possible anymore. The LP (linear programming) and Bayesian approaches of Chapters 4–6 avoid this conflict, as they use benchmark values only as priors that may change in the posterior estimates. A comparison was made of the scope of variables, identities and relations incorporated in the national accounts and in a comparable model of the Curacao economy (Van Tongeren, 2002, p. 325). It was found that

> '... there is more emphasis (in national accounts) than in the model on identities, but national accounts uses less refined methods to extrapolate and only simple coefficients are used to derive the values of some variables from those of others. On the other hand,

> the national accounts, and in particular the 1993 SNA, are much more refined than the model when defining concepts and using sectorization as a means of classifying the data ...'

Vanoli's (2010) comments on the new architecture of the US accounts (Jorgenson, 2008) and his earlier comments on environmental accounting mentioned therein, confirm the separation of accounting and modeling by a prominent national accountant. He accepts 'soft modeling', meaning the use of ratios and identities in conventional national accounts compilation, as long as a sufficient number of basic data items is available, but argues that national accounting should not be extended in the direction of analysis or satellite accounting, if this involves 'hard modeling'.

The compilation approaches of the LP (Linear Programming) and Bayesian SNA (BSNA) frameworks in Chapters 4–6 are very close to conventional compilation practices in national accounting. At the same time is shown that modeling with help of identities and ratios is very much at the heart of national accounts compilation. In a typical national accounts framework of SUT and IEA compiled for a non-base year, only approximately 20% of the variables have basic data (finding based on Guatemala SNA framework of Chapter 6, and excluding price indices as basic data, see below) and the rest is estimated indirectly with help of ratios (assumptions) and identities. The question is, however, whether the 20% scope of basic data mentioned above, which is typical for conventional accounts compilation of SUT and IEA of a non-base year, is or is not sufficient for the use of 'soft modeling' in making the national accounts estimates.

Different aspects of the two questions raised above are dealt with in the chapters of this thesis, and are referred to below. Chapters 2 and 3 deal with the first question of, what is called, the design of SNA and related frameworks. Chapters 4–6 deal with the second question of formalization of the SNA compilation process and the two questions are brought together in the design of the BSNA framework. The latter does not only reflect the data structures of the SNA, but also incorporates other elements that are used in the Bayesian compilation.

2 Design of frameworks in support of analysis

Framework design is a topic that has played an important role in the development of national accounting, and therefore has been dealt with by many authors in the course of time, either explicitly or implicitly, each emphasizing particular aspects. Three main players may be referred to in particular, i.e. the building block approach followed in the US accounts (Vanoli, 2010), the sector-industry approach by Stone as reflected in the subsequent 1968, 1993 and 2008 SNA (UNSD, 1968; United Nations et al., 1993, 2008), and the micro-macro approach promoted by Richard and Nancy Ruggles (1975).

When constructing the first US accounts, a top-down or building block approach was used, in which conceptual links were established between the main macro-economic aggregates identified and measured by Kuznets, i.e. consumption, investments, exports and imports, and GDP. In the course of time those accounts were refined in order to provide users with more detailed information on the macro-economic aggregates. This approach was not fundamentally changed by Jorgenson, when he recently introduced in cooperation with others in the US Bureau of Economic Analysis (BEA) a new architecture for the US accounts (Jorgenson, 2008). What was new in his approach, however, was that he introduced new analytical concepts on productivity and wellbeing into the accounts, which required an extension of the accounts and therefore additional building blocks, including an extension of the constant price aggregates.

Richard Stone, when developing the 1968 SNA and also in his research papers, took a different approach, called by Vanoli (2010) a bottom-up approach, which was thereafter also reflected in the 1993 and 2008 SNA. Instead of emphasizing the macroeconomic aggregates, he emphasized the grouping of economic units, or transactors, into sectors (Government, Non-Financial Corporations, Financial Corporations (banks and insurance), Households, Non-Profit Institutions, and transactions with the Rest of the World) by which transactions of those sectors should be classified; this is referred to as the 'transactor-transaction principle' in national accounting. In doing so, he was able to incorporate input-output tables in the 1968 SNA, by linking data on

ISIC (International Standard Industrial Classification) industry categories in the I-O table to the sector groupings in the national accounts. By emphasizing the groupings of transactors in the accounts, he also paved the way for intensifying the use of classifications in national accounting (see below). In the SNA approach by Stone, the transactions of the sectors in the national accounts were defined in such a manner that macro-economic aggregates were simple aggregates of transactions across sectors.

Richard and Nancy Ruggles had much influence on international standards, partly because Nancy Ruggles occupied a post as senior researcher at UN Statistics Division, where international statistical standards are initiated, and partly because they carried out extensive and critical research of the US accounts. In their studies they emphasized the need to develop micro data sets underlying the national accounts. Those micro data sets would be representative of data on groups of units, which through national accounting would be linked to each other. National accounts with this feature would be an aggregation of the accounts of representative micro units. By maintaining the accounts for micro units, it would be possible to construct different formats of the national accounts serving different types of analysis. Together with their proposals on a micro data base, they also suggested that imputations be limited to a minimum, as statistics should reflect 'facts'. The latter proposal was supported by Van Bochove and Van Tuinen (1986) in a paper prepared for the 1993 SNA revision; it proposed a core system without imputations and specialized modules with imputations. Examples of imputations are the assignment to household final consumption of payments by government and non-profit institutions for education and health, measurement of output of financial services (called FISIM, Financial Intermediation Services Indirectly Measured, in the SNA), that are not recognized by those financial institutions, imputation of government final consumption for the public administration services for which no payments are received, or the imputation of payments for goods and services produced by one and used by another establishment, which are generally not recognized in the business accounts of the enterprise. Strictly speaking even value added may be considered as an imputed concept, which is not recognized at the level of micro units. An argument against the non-inclusion of imputations, is that, while these SNA transactions may not be explicitly taken into account by the units involved, they are essential to arrive at comprehensive concepts for analysis at the macro level.

The framework approach as presented in Chapter 2 and in the BSNA (Bayesian SNA) frameworks of Chapter 6, takes most elements of the above approaches into account, insofar as they are not in conflict with each other. Selected aspects of the frameworks are highlighted in the paragraphs below. It is emphasized that the design of frameworks is a compromise between data availability and analytical requirements.

Frameworks in generic terms are defined as a series of matrices (such as the SUT and IEA matrices) and vectors of variables, between which ratio and identity relations are defined. Basic data are only available for a limited number of variables, and this lack of information is compensated by ratio values and definitions of ratio and identity relations. For the time being only binary ratios are used, but this may change in the future. Frameworks are generally matrices, as they include variables (data plus estimates) that are organized (classified) by sectors, industries, products or other groups of units (columns), and by types of information (rows: SNA transactions, social and demographic phenomena, environmental impacts, etc.). The matrices may also be multi-dimensional, when more characteristics are entered in the definition of frameworks. In order to deal with possible conflicts between the values of basic data and ratios and the identities, prior reliabilities are defined for data and ratio values, which allow for adjustment of these values, so that conflicts are avoided. The identities serve as criteria in arriving at final estimates that are consistent. In the experiences with the Bayesian estimation approach, it was found that the number of basic data plus ratios and identities is generally much larger than the number of variables to be estimated (see Table 1 on page 21), so that the system is overdetermined, and thus checks are available between data and ratio values.

Frameworks are generally defined at the macro- or meso-level and may be aggregations of an underlying data set on micro units, as suggested by Richard and Nancy Ruggles (1975), in which the micro ratio relations are maintained and on which macro consistency identities are imposed. The development of these frameworks would benefit from the use of relational databases (Litwin, 1994), in which data of micro units from surveys and administrative data sources are adjusted (edited and completed), and in which the correspondences between classifications are incorporated (see below).

FIGURE 1

The specific design of frameworks is done in a number of steps. The first step is to define the (policy) issue to be analyzed. Then should be identified the core variables that can be measured potentially and best represent the (policy) issues. And in a third step should be added variables that make explicit the relations between core variables, and thus allow for defining ratio and identity relations between them. This sequence of steps may be illustrated the with help of an example, presented in Figure 1, in support of a study of the interaction between economic development and poverty. The selected core variables representing economic development and poverty may be the well-known indicators of GDP and the number of individuals below the poverty line, as shown in the simple framework. There is no direct relation between the two variables and it is therefore impossible to define any meaningful identities or ratios including both variables that could be used to measure both, if information on only one is available, or to check one against the other, if both types of information are available. The number of information items and variables in this case is the same (2). This changes drastically when incorporating six additional variables:

- socio-demographic variables
  - the number of individuals above the poverty line
  - the population size
  - the (total of) poverty income of those below the poverty line
- the SNA variable of Household Disposable Income that is closest to a measure of wellbeing in the SNA

- variables linking GDP and Household Disposable Income
  - disposable income (=saving) of large corporations
  - net income from abroad.

When using these additional variables, ratios and identities can be defined. The ratios include those between Household Disposable Income and GDP; poverty income and Household Disposable Income; and average Household income per individual below poverty line and average Household income per individual above the poverty line. The identities are those that define the total number of individuals below and above the poverty line as equal to the population of the country, and the identity that defines the relation between GDP and household disposable income.

If one assumes that basic data are available on all SNA variables mentioned, and on the number of individuals below the poverty line and on the population size, the framework has 8 variables to be estimated with 6 basic data, 4 ratios and 2 identities. Thus the system is overdetermined, as there are 12 information items available to estimate 8 variables. The overdetermined system of information can then be used to estimate missing variables and check values of basic data and ratios against each other.

When designing frameworks, priorities are set for the selection of sector, transaction and other detail, in support of the policy issue that is addressed. In the BSNA framework presented in Chapter 6, with emphasis on the statistical needs for fiscal-monetary policies, government and banking sectors and accounts for the rest of the world picture prominently as sectors in the BSNA framework, and also the need to have basic data for those sectors. In a health accounts framework, the NPI and also government sector would show up explicitly, as these are the sectors in which many of the health activities are organized. Furthermore, when designing a framework, one should try to avoid adding together reliable and very unreliable data, so that estimates after reconciliation, based on the reliable data are not affected by estimates based on very unreliable data. For that reasons, data on large banks directly obtained from their accounting records and data on small investment brokers, trusts and other financial auxiliaries often covered through surveys should be covered in separate industries and sectors. Similarly, survey data of small insurance brokers should not be added together with the administrative data of large insurance schemes.

The importance of the framework approach is that data sets are confronted with each other that are often considered separately and thus interactions are brought out between different types of studies that are frequently carried out in isolation. It should be emphasized here, that national accounts and other data frameworks are not the data source for detailed government data, balance of payments data or even NPI data, but rather the instrument through which interactions between those sectors and other parts of the economy are studied. More comprehensive statistics on the sectors mentioned are available from specialized statistical data sources. The fact that national accounts are often limited to GDP estimates is because national accounts are the only data source providing such data. However, using national accounts as a data source and thus developing only GDP estimates, ignores the methodological framework features of national accounts, which has a much wider potential than only providing GDP data.

In the design of frameworks, classifications play an important role. In Chapter 2, where this topic is dealt with, it is recommended that international classifications and correspondences be adapted for national use as follows (see the UNSD website http://unstats.un.org/unsd/methods.htm):

1. All categories of international classifications and correspondences should be dealt with at the most detailed level;

2. All categories of classifications not matching sectors, activities or products should be eliminated and also correspondences between classifications not relevant in the country concerned should be corrected;

3. $M : N$ correspondences between classifications should be converted to $1 : N$ correspondences for pairs of classifications, by splitting up, aggregating or eliminating categories of international classifications. This would apply to the correspondences between sector and ISIC categories, ISIC and CPC (Central Product Classification), COICOP (Classification of Individual COnsumption by Purpose) and CPC; and

4. Classification categories should be aggregated, such that correspondences between classifications are maintained.

These adaptations of international classifications and correspondences facilitate extracting from available data sources a maximum detail of information items that can be confronted with each other in the compilation, so that reconciliation of estimates can take place at a detailed level. With the same aim, the chapter introduces a revised version of the SNA-SUT table in order to integrate more effectively production accounts of industries by ISIC categories, with supply and use balances by CPC categories of products, between the supply of locally produced and imported products and the use of products in intermediate and final demand. In the chapter there is also much emphasis on the use of the present classification by Broad Economic Categories (BEC, UNSD, 2002b), which identifies in foreign trade statistics the economic destinations of products (intermediate consumption, final consumption and capital formation). It is argued that by identifying separately BEC categories in product data that are classified by detailed categories of the product classifications, the supply-use identities are simplified and this improves Bayesian and other estimates. A remaining limitation, however, is that the BEC until now only covers transportable goods in foreign trade statistics, and not non-transportable goods and services. The use of the BEC would be considerably enhanced if it were broadened to non-transportable goods and services categories, so that its limited focus of foreign trade statistics can be widened to other product statistics as well. The classification principles outlined in the chapter were applied in the BSNA frameworks of Chapter 6 to the national accounts data of several countries in Central America.

Other changes in frameworks are introduced to serve analytical purposes. For instance, in the BSNA frameworks used in Central America, two deviations from SNA practices were incorporated. The first one is classification of gross fixed capital formation by industry of use, which permits analysis by industry of marginal capital-output ratios. Another deviation concerns the distinction between imports and local products in all intermediate and final uses of products. The latter distinction was introduced in order to support analysis of the import dependence by products and use. These are by no means the only adaptations to the SNA which could be introduced in SNA-related frameworks. Two important adaptations were for instance proposed in the recent discussion of a new architecture for the US accounts (Jorgenson, 2008). Both referred to constant price estimates. One adaptation was the estimation of capital services in constant prices as part of value added in constant prices. Estimation of this measure together with labour services in constant prices (already in SNA), permitted the derivation of labour and capital productivity measures. Extending constant price measures to saving (=investment in constant prices) in addition to final consumption and investment in constant prices, it was possible to derive a welfare measure, which cannot be derived from the present SNA scope. Another change in economic frameworks is often found in Social Accounting Matrices (SAM's): They include a further breakdown of the household sector by socio-economic categories and a cross-classification of compensation of employees, mixed income and employment of employees and own account workers by ISIC categories and household socio-economic groups, so that the impact of the production structure on poverty in the country can be studied; see Pyatt and Roe (1977) and Thorbecke and Jung (1996).

Changes in the concepts of frameworks are also required when SNA data are reconciled with those of other data systems. This was very much the case in the process which led to the 1993 SNA and more recently to the 2008 SNA, which aimed at harmonizing concepts between the SNA, and the IMF systems of Balance of Payments (IMF, 2010), Government Finance Statistics (GFS) (IMF, 2001) and Money and Banking Statistics (MBS) (IMF, 2000), and ILO's standards on informal sector statistics, which were developed on the basis of resolutions adopted by the International Conference of Labour Statisticians (ICLS). Adjustments were needed as SNA and other systems were often developed in isolation and also with different analytical or policy objectives in mind. A selection of adjustments made to the SNA or the other system may illustrate this. When harmonizing SNA with Balance of Payments standards, the treatment of reinvested earnings as outgoing flows of foreign investors and incoming finance from abroad was adopted in the SNA, and thus national income in the SNA was reduced with the amount of these flows. On the other hand, the SNA distinction between the imports and exports of factor and non-factor services was taken over by the Balance of Payments, as it was an important conceptual element in the definition of GDP. The valuation of imports and exports in the SNA was adjusted to an f.o.b. (free on board) valuation, excluding insurance and transport services incurred during the shipment between countries. Similarly, when GFS standards were harmonized with SNA, the GFS cash basis of payments and receipts was changed to the SNA accrual basis, which is linked to the principles of business accounting. While GFS standards were changed in this respect, it is taking countries a long time, even until now, to adjust to this change in basis of recording transactions in the government sector, as the original GFS concept was very much embedded in country practices. Many changes in SNA concepts took place with regard to the recording and classification of flows and stocks of financial instruments. Particularly during the latest 2008 revision of the SNA, the detail of these instruments was very much refined, and based on MBS concepts, as SNA has much less of a knowledge base in financial operations. Finally, the ILO scope of the informal sector was incorporated as a subsector of the household sector in the 1993 SNA. In the 2008 revision, further details were developed in a separate chapter, in which a distinction was made between the partly overlapping non-observed economy and the informal sector, sectorization of the SNA between the household and non-financial corporate sector was refined with regard to the classification of informal sector activities, and furthermore a distinction was made between the not fully overlapping categories of informal sector production activities and informal employment.

Concepts are even more subject to change when SNA is extended to satellite accounting, as often non-monetary variables are incorporated in the framework and also analyses alternative to those of the SNA are implied by those frameworks. Chapter 3 on environmental accounting presents a first and important example of this. In it the essence of the SNA framework is maintained, while at the same time identifying separately activities producing natural growth products in agriculture, forestry and fishing and environmental protection services, and including a so-called eco-margin as additional cost of production, which lowers net value added. Net value added and Net Domestic

Product are important concepts in environmental accounts. Environmentally adjusted value added adds up to Environmentally adjusted Domestic Product (EDP), which is an alternative to NDP and GDP in SNA. The difference between NDP and EDP is the eco-margin, which includes the imputed market value of maintenance, avoidance or restoration cost of emissions and degradation of land, water and air and also the cost of depletion of e.g. subsoil resources or wood in forests. Correspondingly, the asset boundary of the balance sheets in SNA is extended to all natural produced and non-produced assets with a market value, either because the assets are regularly depleted for market purposes (e.g. ocean fish, tropical wood and subsoil assets) or because the assets are directly marketed (uncultivated land in exceptional cases). The environmental indicators are incorporated in an extended SNA framework, which combines the use of the SUT with asset accounts and balance sheets (Table 1 in Chapter 3). In this table the eco-margin is incorporated as an adjustment of the SNA concept of capital accumulation of produced manmade and natural assets and non-produced natural assets which, together with revaluation of assets, measures the difference between opening and closing stocks of assets. As all flows and stocks in the environmental accounts framework are valued in monetary terms, the paper deals extensively with the valuation issue of depletion and degradation and stocks of natural assets. In Ward (2004) this development is considered as the very first effort to integrate economic and environmental data in one framework, and was the basis for the first System of Economic-Environmental Accounts, SEEA (UNSD, 1993). Since then many other environmental SNA extensions have been proposed and/or implemented, most of which have been reflected in the latest version of the SEEA (United Nations et al., 2003). The paper furthermore sets the stage for a discussion of the concept of 'sustainability', which it considers as a cost concept as distinct from a wellbeing concept, which uses the 'willingness-to-pay principle' as a valuation concept. As valuation remains a controversial topic until now, the 2003 SEEA includes a so-called 'hybrid' SUT. In this table SUT data in monetary units on value added and products used in intermediate and final demand are combined with data in physical units on the use of natural resources in intermediate and final use, and on degradations due to residuals and wastes affecting produced and non-produced assets. The main concepts of EDP and capital accumulation are maintained in the 2003 SNA, but more precisely defined. The term eco-margin is not used anymore.

Health accounts are another satellite system that has been developed internationally (OECD, 2000; PAHO, 2005). It extends the SNA scope in order to juxtapose, for purposes of analysis, data on the occurrence and treatment of diseases with data on the production and consumption of health services combating those diseases, and data on the finance of health services by social security and private insurance schemes. In order to achieve this integration of health and economic and financial data in one framework, and thus develop an effective instrument for health policy, many adaptations are needed to integrate health-related classifications of diseases and treatments of diseases, with ISIC classifications of production units (producers of health services), functional classifications of household final consumption (of health services) and sector classifications on finance of health services used in the SNA; see Van Tongeren (2006). Emphasis in the integration of health accounts and SNA is on the SNA concept of 'actual final consumption' of health services, which includes not only final consumption of those services paid for by households, but also expenditures on health services paid for by government and non-profit institutions, in order to provide a comprehensive picture of all expenses on health services. The discussion on health accounts also deals with the extension of the SNA production boundary, in order to encompass not only market support of health, but also family and other volunteer support of health services, which are not included in the market concept of GDP. When extending the SNA to health accounts, another issue arises, i.e. the use of partial information, as practiced in SNA accounting, versus the use of comprehensive data sets in health accounts. The use of partial data sets in SNA compilation leads to much saving in time and labour resources, and as is shown in Chapter 6, the precision of the estimates is not affected in any major way. As the use of complete data sets in health accounts is very labour and time consuming, a similar strategy as in national accounts may be adopted in this satellite extension as well.

The application of the Bayesian methodology is not limited to the issues, on which already frameworks have been defined. They can be extended to other uses, as was shown in the example of Figure 1. In this context may be mentioned several applications that have been started or are in the planning stage. A regional GDP application has been carried out in the Kurdistan region of Iraq. Quarterly accounts in the Dominican Republic are presently being prepared as an extension of the BSNA in Central America. In the same context proposals have been made to develop projections for those countries, in which the impacts of policy changes can be assessed and also projections for future periods are being considered. The advantage of the latter is that the projections follow the same methodology as compilations for past periods, and that estimates resulting from the Bayesian approach do not have to be inserted in separate models for projection to the future. Another application that is already in advanced stage of compilation is an economic-environmental-health satellite framework for a region in Guatemala, in which the health impacts of economic activities are measured through the environmental medium of water in a lake, the quality of which affects of the health of the population that is using the water. The framework approach may also be developed in the future as a preferred alternative to aggregate indicators, such as the Human Development Indicator (HDI) of UNDP, in which per capita GDP is merged with indicators of life expectancy and education, or the National Happiness indicator of Bhutan, which was developed on the basis of recommendations by Stiglitz et al. (2009). In those instances, the indicators merged with GDP may instead be brought together in a framework and linked as was done in Table 1, so that double counting is avoided.

3 Formalization of compilation

Chapters 4, 5, and 6 reflect the development over time of the formalization of SNA compilation methodology. The paper on which Chapter 4 is based was one of the first responses to the question of formalization of the compilation process. It uses a semi-fictitious 1965 data set of Surinam, supporting a small set of 48 variables without an explicit framework, covering SUT and IEA aggregate data on the main sectors of the economy. Among the ratio values included in the framework are growth rates of variables between two periods. It defines prior reliabilities of data and ratio values as discrete intervals, and makes posterior estimates of variables and ratios, taking into account the prior intervals and national accounts identities that should hold between the variables. An LP approach was used to arrive at posterior intervals and central values of the estimates that would satisfy all conditions. Chapter 5 presents an explicit framework, based on a fictitious data set, and using a small set of 38 variables, similar to that in Chapter 4, including some growth rates of major variables such as GDP. The discrete prior intervals of data and ratio values have been replaced by continuous intervals, using the assumption that those values have a normal distribution with prior coefficients of variation that are set at values based on national accountants' assessments. Thus, in the latter methodology, posterior values (= estimates) of variables and ratios could, in principle, take all values between $-\infty$ and $+\infty$, while in the LP methodology of Chapter 4, the posterior estimates can only be located within the prior intervals that were defined. Chapter 6 applies a similar methodology as in Chapter 5, but increases the number of variables from the 38 (Chapter 5) to 2719 (Chapter 6). The BSNA framework of Chapter 6 includes detailed SUT and IEA data sets, using detailed classifications of data by sectors, industries (ISIC) and products (CPC). Also included/identified in the framework are the basic data, ratio definitions and values, identity definitions and reliabilities (see Figure 1 in Chapter 6). Prior to making final Bayesian estimates in this chapter, tentative estimates are made using a selection of identities and ratios (see Section 6.3). In order to deal computationally with this large data set, so-called 'sparse matrices' of data and restrictions are used, from which the large number of zero elements, have been eliminated.

The three frameworks differ in the extent to which prior conditions affect the posterior estimates. This is reflected in Table 1, which compares the number of variables, basic data, ratios and identities between the schemes. Two assessment indicators are derived: The first one is the percentage of variables supported by basic data, which is the highest (75%) in the case of the LP framework of Chapter 4. The percentage of basic data thus derived

Table 1: Comparison of frameworks and estimates

	Framework		
	Ch 4	Ch 5	Ch 6 (2006YF)
Nr of variables	48	38	2719
Nr of basic data	36	19	531
Nr of identities	4	16	1120
Nr of ratios and other priors (incl. price indices)	13	36	2294
Information factor	1.10	1.87	1.45
% of variables supported by basic data	75.0	50.0	19.5

for the BSNA framework (19.5%) and quoted in Chapter 6 is relatively low. However, this does not take into account as basic data, the price indices, which are also directly measured; they are treated as ratios. If treated as basic data, the percentage increases to close to 50%. The second indicator is the information factor, which divides the number of information items (basic data, ratios and identities) by the number of variables to be estimated. The more information items, and particularly the more basic data, are available, the more precise will be the estimates. The information factor is the highest (1.87) for the framework of Chapter 5, the lowest (1.10) for the LP framework of Chapter 4 and in-between (1.45) for the Bayesian framework of Chapter 6.

All three chapters include sensitivity analyses of the impacts of changes on posterior estimates of data availability and other changes in the configuration of the frameworks. In Chapter 4 it is shown that if no data are available on imports and production accounts data of other industries (mainly services), the posterior intervals of the estimates are considerably widened. In Chapter 5 it is shown that a 'reduced' compilation, in which the number of basic data is only half of the number in a 'comprehensive' compilation, estimates generally deviate much more from conventional estimates that are also simulated in the chapter. Chapter 6 includes various tests of the BSNA framework, including the impact of fewer data annually available, different prior reliabilities, use of 'tentative' estimates, different scopes of the framework, different availability of basic survey and administrative data, and of aggregation of basic data. In the sensitivity analyses of all three chapters, estimates are compared with conventional national accounts estimates. The reason behind this is that for the time being, the Bayesian (incl. LP) methodologies are trying to simulate as best as possible the conventional compilation methods (using the same basic data, ratios and identities, and making explicit prior reliabilities) and arrive at estimates that are as close as possible to the conventional estimates. This is also the criterion by which the methods are being judged by users, particularly if the Bayesian compilation methods are used alongside the conventional methods, until the point is reached that they can fully replace the conventional methodology. At that point it is possible to change the orientation and experiment with the Bayesian methodologies aiming at improving the conventional estimates.

FIGURE 2

An important point in the development of a formalized SNA compilation methodology was the publication of 'A Systems Approach to National Accounts Compilation' (UNSD, 1999). In this document the process of conventional national accounts compilation and integration was described, as a sequence of steps between basic micro data and reconciled SNA aggregates at the macro level. The essence of the approach is summarized in Figure 2. It includes the creation of four different data sets from micro data to reconciled national accounts data, which are linked to each other through three compilation steps. The preparation of sector and industry data in national accounts format, after micro data have been edited, aggregated and converted, is the point at which data are available in non-reconciled format. The data in these accounts are brought together in the SUT and IEA, which bring out the discrepancies between the data and are also used to reconcile those.

For each of the steps, tables were suggested, in which the data were converted from the format of the previous step and prepared for the next step in the compilation. All tables included in the publication are supporting the standard framework of the international SNA. All steps, except the last reconciliation step, could be supported by simple computer processing: This applies to the first compilation step, in which micro data from surveys and administrative data sets after editing are aggregated to groups of micro units. Those are a combination of ISIC and sector categories used in the last step of reconciliation and groups of units that are separately handled in the organizational set-up of a national accounts department and which reflect very much the statistical sources of basic data that are highlighted at the lower end of the table. Conversion to national accounts format in the next compilation step is also simple in terms of computer processing. Computerization of the last step is different, however, and more difficult as it involves a confrontation of different values directly or indirectly estimated for the same SNA variables, which even until now many still prefer to reconcile manually (see Bos, 2009b), as they involve assessment and elimination of discrepancies and decisions about which data to adjust and how, in order to achieve this.

The formalization and thereafter computerization of the last step has been introduced in two different ways in practical applications. One way is to simulate the sequential reconciliation process as done conventionally by national accountants and provide users with the opportunity to take decisions at different stages in the compilation process. This is done in one of the widely-known software programmes written for national accounts compilation, i.e. ERETES (1997). It focuses first on the reconciliation of the supply-use balances of the SUT, and then reconciles the resulting estimates within the constraints of the receipt and expenditure balances between sectors of the IEA. In this process data are being adjusted to satisfy the supply-use balances of SUT and the receipt and expenditure balances between sectors of the IEA. In this sequential reconciliation process of the conventional methods and ERETES, priority is given to the SUT estimates, which are made first, over the IEA estimates, which are made after SUT estimates have been reconciled. IEA restrictions on the SUT estimates are not fully taken into account in these approaches. This limitation of the conventional and computerized ERETES methods is overcome in the simultaneous reconciliation of the LP and Bayesian estimation methods as described in Chapters 4, 5 and 6. All values of data and ratios are simultaneously reconciled, under conditions of identities, ratio definitions and prior reliabilities, and thus IEA restrictions of data, ratios, identities and reliabilities may also influence the estimates of the variables in the SUT.

Another question that is raised, is whether in the different compilation approaches always solutions can be generated, and if so, whether those are unique solutions? The sequential conventional and computerized methods (e.g. ERETES) always have solutions, as any incompatibilities between data and assumptions are overcome by revising those in the process of sequential compilation. Also, they are not unique solutions, as they are very much dependent on subjective decisions of the compiler at different stages of the compilation process; another compiler with different views may arrive at different estimates.

The discrete LP method of Chapter 4 does not always have a solution. This may become clear from Diagram A in Chapter 4, which presents the LP method graphically. When using the so-called Simplex method to arrive at posterior intervals (maximum and minim values) of variables under conditions of identities and maxima and minima of ratio values and basic data, solution vectors are obtained which include values of all variables, some of which are maxima or minima. The central value of each variable is then obtained as part of a solution vector that is the unweighted average of the extreme solution vectors. As final estimates in this method cannot lie outside the prior ranges of data and ratio values as indicated in the diagram, and if those ranges are incompatible with the identities, there may not a feasible solution under all circumstances, using this method. Due to the small size of the framework in Chapter 4, this problem was not a serious one, as prior ranges of selected basic data or ratios could be easily adjusted in order to overcome the problem of no solution. It is expected, however, becoming a very serious problem, indeed, if the framework would become as large as the one that is described in Chapter 6. It would lead to many intervals that are confronted directly or indirectly with each other, and such large number of confrontations may be unmanageable in practice.

The solution issue was largely overcome by using the continuous reliability intervals of the normal distribution in Chapters 5 and 6. In that approach it is possible to arrive at posterior estimates (central values) for all variables and ratios. The Bayesian estimates are unique estimates; subsequent compilations under the same conditions would generate the same estimates. Bayesian estimates would also be generated, in those cases that no solution would be obtained through the LP approach. It is expected, however, that if no solutions could be found in the discreet LP approach, the Bayesian approach would generate 'nonsense' solutions. Unfortunately, no comparable estimates of the discreet LP and continuous Bayesian approaches are available at present for large data sets, so that it is not yet possible to confirm this intuitive statement.

Reliability remains a concept in the LP and BSNA frameworks that still needs further development. Prior reliabilities are implicit in the sequential conventional compilation methods and are brought out explicitly in the prior intervals of the LP method and in the prior coefficients of variation in the Bayesian estimation based on normal distribution of prior values. The prior reliabilities remain subjective, however, in the LP and Bayesian methods. The subjectivity varies between the prior information that is available: Prior reliability of basic data from surveys may be determined with help of sampling methods. Reliability of basic data from administrative data sources, such as government, banking, and foreign trade records can only be subjectively determined, and such subjective assessment should take into account that such records cannot be changed, even if statisticians have their doubt about the quality of those data. Assessing the reliability of ratio values should take into account that the basic data of the variables underlying the ratios are often derived from base year data that are unreliable themselves, and also the volatility of those ratios over time, even of I-O technical coefficients. However, since Morgenstern (1970) for the first time raised the issue of reliability of economic statistics and national accounts, much work has been done internationally, with the help of IMF's Data Dissemination Standards (DDS) and Quality Assessment Frameworks (DQAF), to assess the reliability of statistics in general and national accounts in particular, not only focusing on the accuracy of the statistical information, but also dealing with accuracy and reliability of source data, accessibility of data, timeliness of data, methodological soundness of the data compilation, professional integrity of the compilers, etc. These efforts very much contribute to determining the prior reliabilities of single items used in the Bayesian estimation approach. The LP and Bayesian approaches, however, introduce the concept of posterior reliabilities of data and ratio values, and those reliabilities do not refer to single items, but rather to a set of posterior reliabilities that are derived as a result of making estimates in a framework.

References

Bos, F. (2009a). *The National Accounts as a Tool for Analysis and Policy in View of History, Economic Theory and Data Compilation Issues*, VDM Verlag Dr. Müller, Saarbrücken.

Bos, F. (2009b). The art and craft of compiling national accounts statistics and their implications for reliability, *The Review of Income and Wealth*, 55, 930–958.

ERETES (1997). National accounts software, INSEE (France) and EUROSTAT (http://www.eretes.net).

IMF Statistics Department (2000). *Monetary and Financial Statistics Manual*, Washington, DC.

IMF Statistics Department (2001). *Government Finance Statistics Manual*, Washington, DC.

IMF Statistics Department (2010). *Balance of Payments and International Investment Position Manual*, Washington, DC.

Jorgenson, D.W. (2008). A new architecture for the U.S. national accounts, Richard and Nancy Ruggles Memorial Lecture, 30th General Conference of the International Association for Research in Income and Wealth, Portoroz, Slovenia.

Litwin, P. (1994). *Fundamentals of Relational Database Design*, adapted by K. Getz, P. Litwin, and G. Reddick from Microsoft Access 2 Developer's Handbook, Sybex (http://www.deeptraining.com/litwin/dbdesign/ FundamentalsOfRelationalDatabaseDesign.aspx).

Morgenstern, O. (1970). *On the Accuracy of Economic Observations*, Princeton, pp. 242–276.

Organisation for Economic Cooperation and Development (OECD) (2000). *A System of Health Accounts*, Version 1.0, Paris.

Pan American Health Organization (PAHO) (2005). Satellite Health Account (SHA), Manual, Version 1, Health Policies and Systems Unit (HP), Manual prepared by M. Ascues, M. Castro, and C. Reyes, with the collaboration of A. Gordillo.

Pyatt, G. and A. Roe (1977). *Social Accounting for Development Planning with Special Reference to Sri Lanka*, Cambridge University Press, Cambridge.

Ruggles, R. and N. Ruggles (1975). The role of micro data in the national economic and social accounts, *The Review of Income and Wealth*, 21, 203–216.

Stiglitz, J.E., A. Sen, and J.-P. Fitoussie (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress* (www.stiglitz-sen-toussi.fr).

Stone, R. (1955). Model building and the social accounts, a survey, *The Review of Income and Wealth*, 4, S27–S77.

Stone, R. (1984). *The Accounts of Society*, Nobel Memorial Lecture.

Thorbecke, E. and H.-S. Jung (1996). A multiplier decomposition method to analyze poverty alleviation, *Journal of Development Economics*, 48, 279–300.

United Nations Statistics Division (1968). *A System of National Accounts, Studies in Methods*, Series F, No. 2, Rev. 3, UN, New York.

United Nations Statistics Division (1993). *Handbook of National Accounting-Integrated Environment and Economic Accounting* (Final draft circulated for information prior to official editing), Series F, No. 61, UN, New York.

United Nations Statistics Division (1999). *A Systems Approach to National Accounts Compilation*, Studies in Methods Series F, No. 77, Handbook of National Accounting, Sales No. E.99.XVII.10, UN, New York.

United Nations Statistics Division (2002a). *Handbook on National Accounting: Use of Macro Accounts in Policy Analysis*, National Accounts Handbook Series, UN, New York.

United Nations Statistics Division (2002b). *Classification by Broad Economic Categories, Statistical Papers*, Series M, No. 53, Rev. 4, UN, New York.

United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, and World Bank (1993). *System of National Accounts, 1993*, United Nations publication, Sales No. E.94.XVII.4.

United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, and World Bank (2003). *Handbook of National Accounting—Integrated Environment and Economic Accounting*, Series F, No. 61, Rev. 1.

United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, and World Bank (2008). *System of National Accounts, 2008*, United Nations publication, Sales No. E.08.XVII.29.

Van Bochove, C.A. and H.K. van Tuinen (1986). Flexibility in the next SNA: The case for an institutional core, *The Review of Income and Wealth*, 32, 127–154.

Vanoli, A. (2010). The new architecture of the U.S. national accounts and its relationship to the SNA, *The Review of Income and Wealth*, 56, S734–S751.

Van Tongeren, J.W. (2000). Concepts, measurement and analysis of macro accounts, paper presented at the ‘System Of National Accounts 1993: Unfinished Businesses’ session of the Annual Meeting of the American Economic Association, Boston.

Van Tongeren, J.W. (2002). Integration of accounting and modeling in the Caribbean: The Antillyse model and the national accounts of the Netherlands Antilles, in: United Nations Statistics Division (2002a, pp. 315–331).

Van Tongeren, J.W. (2006). Review of OECD Health Accounts in relation to SNA, paper prepared for OECD meeting on Health Accounts, Paris, 5–6 October 2006.

Ward, M. (2004). *Quantifying the World: UN Ideas and Statistics*, UN Intellectual History Project Series, Indiana University Press.

Figure 1 Poverty-economic development frameworks

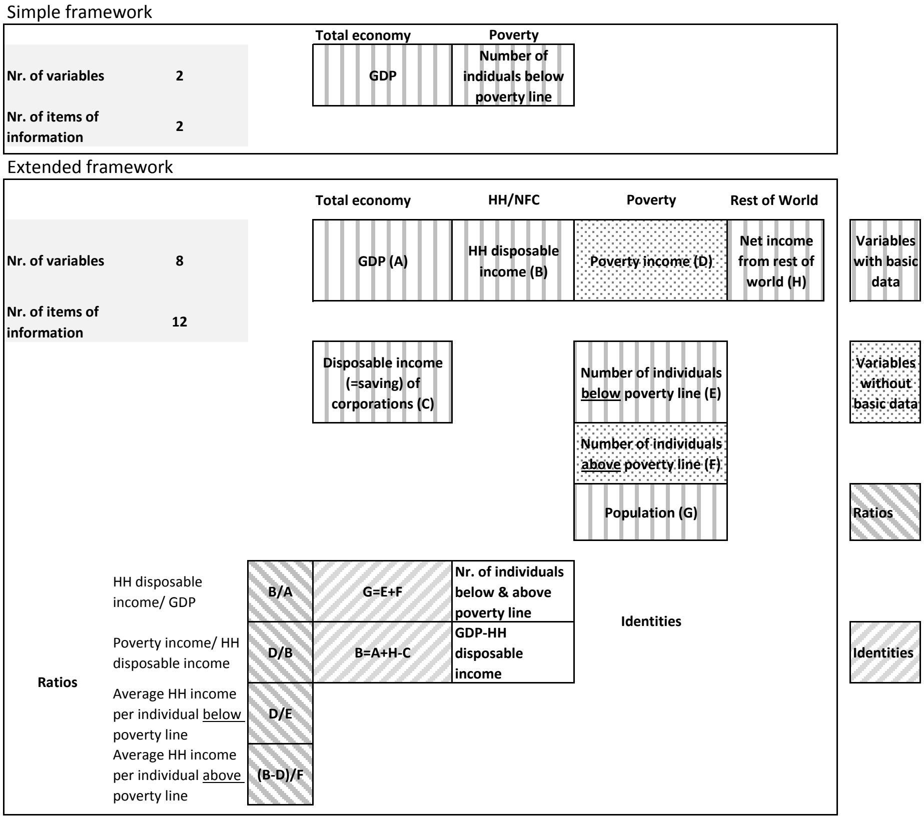


Figure 2 Elements of Systems Approach to compilation of national accounts

Reconciled national accounts data

↑

Reconciliation through further adjustments of intermediate data in the national

↑

Sector and industry data in national accounts format

↑

Adjustment of intermediate data and their conversion to the national accounts

↑

Intermediate data after editing and aggregation

↑

Editing and aggregation of basic data

↑

Micro data obtained from surveys, censuses and administrative records

Economic surveys and censuses, incl. surveys of small scale industries and surveys of non-profit institutions.
Agricultural surveys and censuses

Family income and expenditure surveys
Labour and occupational surveys
Demographic surveys and censuses

Banking and insurance statistics prepared by Central Bank, insurance and other institutions supervising the

Financial statements of individual private non-financial corporations, consolidated statements of supervisory institutions, tax records , surveys of enterprises

Government budgetary records and financial statements of public corporations and decentralized public administrative bodies, also including government finance statistics.

Financial statements of individual private non-profit institutions, surveys of private non-profit institutions

Balance of payments statistics, largely based on Central bank foreign exchange records
Foreign trade statistics based on customs' records

Chapter 2

Designing analytical data frameworks

Review of Income and Wealth
Series 50, Number 2, June 2004

DESIGNING ANALYTICAL DATA FRAMEWORKS

BY JAN W. VAN TONGEREN*

Development Research Institute (IVO) of Tilburg University, Netherlands

The article is about design of analytical data frameworks that make optimal use of micro data available to serve the needs of the policy analyses and create a maximum number of data checks at the macro or meso level. As the SNA is one of the most extensive analytical data frameworks, the notions developed in this paper will be applied mainly to data segments of the SNA and in particular to a much used data segment of the SNA, i.e. the Supply and Use Table (SUT). The larger part of the paper focuses on defining a simplified SUT and using a so-called "system of classifications and correspondences," that is based on international SNA and classification standards but taking into account specific features of the economy of a country, data availability and the type of analysis served by the framework. The simplified SNA should be comprehensive, but should avoid unnecessary cross-classifications of data and details.

INTRODUCTION

Designing data frameworks, even within the limited context of the SNA, is a much-neglected topic in which very few authors have expressed interest. Ruggles and Ruggles were among the few who did see the importance of this topic in relation to the SNA (Richard and Nancy D. Ruggles, 1975; Nancy D. and Richard Ruggles, 1983, 1986, 1992).

The objective of the present paper is to develop guidelines for the design of analytical data frameworks, which have the following features:

- Available data obtained from micro units are incorporated in the framework as aggregate data sets for groupings of micro units, referred to here as meso groupings or sectors. The meso groupings or sectors selected are those that are the subject of government and other policies.
- When aggregating the micro data into data sets for meso groupings,
  - a maximum amount of analytical information should be obtained from the available micro data;
  - a maximum amount of data checks available in the micro data should be incorporated, so that, after adjustment of data in line with the consistency requirements of data checks, optimal data consistency and reliability is achieved.

When designing an analytical framework, a delicate balance should be kept between comprehensiveness in terms of analytical information and data checks

Note: The author wishes to thank his former colleagues at the UN Statistics Division (Ralph Becker, Ivo Havinga and Vu Viet), and his present colleagues at Tilburg University (Arthur Giesberts, Judith Lammers, Jan Magnus, Ruud Picavet) for the technical support and constructive comments received on earlier drafts of this paper.

*Correspondence to: Jan van Tongeren, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands (JanWvT@compuserve.com).

that are included, and simplification and transparency of the framework in which unnecessary details are avoided. The latter is emphasized in Section 2.3, when referring to a simplified SNA.

In the remainder of this paper, references will be mainly to the 1993 SNA, as it is the most comprehensive representative of framework development. If its most extensive core of Supply and Use Tables (SUT) and Integrated Economic Accounts (IEA) were implemented, it would be able to enforce data consistency and analysis based thereon for almost all economic statistics; and if extended to satellite accounting, it is possible to extend such data consistency to social and environmental statistics, which can be linked to economic data. Frameworks also exist in specialized fields of socio-economic statistics, including IMF frameworks for Government Finance Statistics (IMF, 2001), Balance of Payments (IMF, 1993) and Monetary and Financial Statistics (IMF, 2000), Labor Accounts (Hoffmann, 2000) and Informal Sector statistics (ILO, 2001), developed by the ILO or frameworks for Agricultural Accounts and Food Balances developed by FAO (FAO, 1996). For population projections there are frameworks defining the relation between population variables, such as birth, death, and migration and between the changes in those variables (United Nations, 1983a).

Tradition in national accounting, and possibly in other frameworks, has it that tables in which data are presented are determined by international and regional SNA standards. While this practice contributes to international comparability, it makes less than optimal use of the available data when incorporating checks in the data compilation and when revealing the analytical features of an economy. In the present paper it is suggested that the international standards of data presentation are supplemented by analytical frameworks that are specifically designed for local use, in support of effective local analysis and compilation of data.[1] When describing the design of the SNA framework for local use, the paper uses a formalized description of the SNA that differs in some respects from that of the 1993 SNA Manual, in which reference is made to terms such as data vectors, matrices, identities and indicator ratios, systems of classifications, attributes, etc., which are not part of the terminology used in the SNA Manual.[2] They are, however, essential in the formalized description of the System pursued here.

[1]The topic of this paper is part of a continuing research project at Tilburg University carried out by the author and others to formalize a number of aspects of SNA and other analytical frameworks that are determined mainly by tradition until now. The research focuses on three aspects, i.e. (i) the design of frameworks, (ii) the approach to compiling the framework data (see van Tongeren, 1986; Magnus, van Tongeren, and de Vos, 2000), and (iii) the relation between the assumptions used in compiling framework estimates and analytical indicators in which there is much international interest at present. The intended formalization of national accounts and other frameworks is needed in order to make more transparent the process of designing those frameworks and developing framework estimates. It furthermore is an essential requirement of the increasing computerization of the national accounts compilation process. The present paper, while focusing mainly on the design of frameworks, will sometimes make reference to compilation and indicator ratio issues, but will not deal with these issues in detail.

[2]The term "attribute" is used in databases to indicate the various dimensions of information items included in the database. The term is used in the present paper in relation to classifications and cross-classifications, and implies that the suggested approach to classifications is close to a database approach and/or may be adapted to it.

1. Analytical Data Frameworks

The term analytical data framework is being used frequently by many, but in a loose manner. As the term is used throughout this paper, there is a need to define it more precisely.

An analytical framework is a table and/or a group of tables, which include data, and with relations defined between the data of those tables or groups of tables. The relations between the data included in a framework may be used in analysis (this explains the adjective "analytical"), and some of the relations may also be used in the compilation when checking the consistency of data in the tables. The tables may be one-dimensional vectors, or two- or multi-dimensional matrices. The relations between the elements of the tables may be identities or indicator ratios. The only relation that does not qualify a table to be an analytical framework is the identity between the row and sum-totals, which always holds in every table. The relations of identities and indicator ratios defined between the data of the framework are not only used in the analysis of the data, but also define the consistency of the data included in the framework.

An analytical data framework differs from a database. The latter, supported by several computer software programs, only organizes the data, using various attributes that are defined for each data element. A database, however, does not necessarily include relations between the data, as is the case in analytical data frameworks.

The relation between meso aggregation, data checks, and analytical information is reflected in the illustrative scheme of Table 1, which is based on the 1993 SNA. The data of meso groupings or sectors shown in the columns refer to the resident sectors that are distinguished in the SNA and that interact with each other in the economy of a country; they are the Non-Financial Corporations (NFC), Financial Corporations (FC), Government (GOV), Households (HHs) and HH subsectors (such as urban and rural HHs, or HHs with different composition and income or other characteristics) and Non-Profit Institutions (NPI).[3] The information items in the columns of the resident sectors are aggregated into those of the National Economy. The column for the Rest of the World (ROW) includes counterpart items such as exports and imports that constitute economic and other cross-frontier impacts of actions of resident sectors outside the economy and impacts from outside the economy on resident sectors. The sector columns are augmented with a column for the grouping of natural resource units (forests, fish stocks, mineral reserves, quality of air, etc.) that is used in economic-environmental satellite accounts analysis and brings together the information items on impacts of the other sectors on the natural environment.

In the column of each sector are included data boxes for economic, social and environmental information items that describe the actions and impacts thereof for each sector.[4] The upper panels of economic data items refer to those of the

[3]It was thought to be more appropriate to use in the generic scheme of Table 1 the more general term of NPIs (Non-Profit Institutions) rather than the more specific SNA concept of NPISH (Non-Profit Institutions Serving Households).

[4]The 1993 SNA uses for the majority of the information items the terms "transactions, other flows, and stocks" (see 1993 SNA sections, B, C and D of Chapter III). As these terms, however, are too narrow for information items that relate to economic-environmental and socio-economic accounting, it has been preferred here to use the more general term "information item."

TABLE 1

Analytical Framework of Comprehensive SNA, Including Economic-Environmental and Socio-Economic Satellite Accounts

Aggregates of meso groupings

Meso groupings

Total Economy	ROW	NFC	FC	GOV	HH and HH Subsectors	NPI	Natural Resources by Type	Horizontal identities data balances (=checks)
Output Intermediate consumption Capital formation Final consumption	Exports Imports	Output Intermediate consumption Capital formation	Output Intermediate consumption Capital formation	Output Intermediate consumption Capital formation Final consumption	Output Intermediate consumption Capital formation Final consumption	Output Intermediate consumption Capital formation Final consumption	SUT	Supply use identities
Value added and components		Value added and components	Value added and components	Value added and components	Value added and components	Value added and components		Resource use identities
Compensation of employees Property income Current and capital transfers	Compensation of employees Property income Current and capital transfers	Property income Current and capital transfers	Property income Current and capital transfers	Property income Current and capital transfers	Compensation of employees Property income Current and capital transfers	Property income Current and capital transfers		
Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes		Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes	Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes	Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes	Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes	Produced and non-produced non-financial assets: stock/ purchases/value and other volume changes	Natural assets: opening and closing stock (in quantity and quality units), other volume changes	Balance sheet identities
Financial assets: stock/ acquisition/value and other volume changes	Financial assets: stock of external assets held by residents, acquisition of external assets by residents	Financial assets: stock/ acquisition/value and other volume changes	Financial assets: stock/ acquisition/value and other volume changes	Financial assets: stock/ acquisition/value and other volume changes	Financial assets: stock/ acquisition/ value and other volume changes	Financial assets: stock/ acquisition/value and other volume changes		
Liabilities: stock/ acquisition/value and other volume changes	Liabilities: Stock of external liabilities, incurrence of external liabilities by residents	Liabilities: stock/ acquisition/value and other volume changes	Liabilities: stock/ acquisition/value and other volume changes	Liabilities: stock/ acquisition/value and other volume changes	Liabilities: stock/ acquisition/value and other volume changes	Liabilities: stock/ acquisition/value and other volume changes		
Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Emissions into air, water and land	Material balances
Employment population	Employment of migrant workers	Employment	Employment	Employment	Employment Population groups	Employment		Population vintages

Vertical Identities and indicator ratios defining 'behaviour' of meso groupings economic core of the SNA, while environmental and social data refer to information items in environmental-economic (United Nations, 2003) and socio-economic satellite accounts (National Statistical Office, Republic of Korea, 2000) and Social Accounting Matrices (see Thorbecke (2000) for the latest overview of the work on SAMs). The information items included in each meso sector column are obtained by aggregating the corresponding information of micro units. Illustrative examples of information items are included in the presentation of the scheme for each sector.

Requirements of analysis and policy use determine the selection and detail of sectors and the selection of the information items to be included for each. Indicator ratios between values of the information items in each column, such as input-output coefficients, labor productivity measures, tax ratios, etc., represent the instruments of analysis embedded in the scheme; and the values of some of these indicator ratios may also be used as data checks, when verifying the mutual consistency of the values of the information items. Other data checks are horizontal identities that may be defined between information items of different sectors. On the right-hand side of the table are included five types of horizontal balances that should hold in a comprehensive SNA, including (i) balances of supply and use of products of the SUT, (ii) resource–use identities in the IEA, that apply to receipts and disbursements of property income, current an capital transfers, etc. between sectors, and (iii) balance sheet identities between stocks of assets of sectors and financial claims on those assets by other sectors. (iv) The material balances of environmental accounting show how emissions "caused" by sectors have an impact on (or are "borne by") the natural resources in the last column, and (v) the identities of population vintages show how population groups move in the course of time from schools to employment, unemployment and retirement, etc. When more detail is included on the information items of each sector, more horizontal data checks and also more indicator ratios become available.

2. Adapting and Further Detailing the SNA-SUT with the Help of Classifications and Correspondences

An aggregate presentation of the economy and its social and environmental ramifications is useful for quantification and analysis at the macro level of concepts, such as production, investment and consumption, or impacts on the environment and population and employment groups. It is, however, too macro-oriented to identify and measure data structures describing behavior and other characteristics of groups of agents that are the target of economic, social and environmental policies, such as producers, households and groups of individuals including those employed, social institutions (schools, hospitals), natural resources, etc. Maintaining the overall macro structure, while incorporating more detail on meso sectors and data for each sector, is what is often called maintaining the requirement of micro-macro links.

The detail of meso sectors is implemented through the use of classifications. As many data sets are joined in a framework, many classifications are needed and several data sets can be classified according to different criteria, using binary or even multiple cross-classifications of the same data. The cross-classifications are an important instrument of analysis, as they establish relations between the data details from different points of view, which is what is most useful when doing studies in support of policy analysis.

How classifications and correspondences will be used in the further detailing of an analytical framework will be illustrated in the remaining part of this article with help of a subset of the comprehensive SNA in Table 1, i.e. the SUT (Supply and Use Table). It is an SNA segment identified in the upper part of Table 1 and the variables included are located within the solid lines marked there. The SUT segment was selected because it combines several features that are relevant to analytical frameworks: i.e. many analytical relations are defined, and most SNA classifications interact in this part of the System. Furthermore, the application of framework features to the SUT is relevant as there is extensive use of this type of information in I-O types of analysis, and the SUT is also the starting point for several types of satellite accounts, including socio-economic, environmental-economic, tourism, export satellites, etc.

2.1. *International Classifications and Correspondences*[5]

Adaptation of classifications and cross-classifications should start from what is internationally available. The extent to which international classifications and correspondences may be used in specific country settings, however, depends on the specific features of an economy, the analytical objectives of the accounts, and also on data availability. In order to use the international classifications and correspondences in the most optimal manner for detailing analytical frameworks in general and the SUT in particular, the user is provided below with an overview of what classification detail is internationally available at the most detailed level and what type of correspondences are defined between them.[6] This information is generally not available from the internal classification references referred to above.

- The most detailed classification is the HS with 5224 categories at the most detailed level, and restricted to goods only.[7] The CPC (version 1.0) has fewer categories at the detailed level, covering both goods and services (CPC version 1.0: 2101 and CPC 1.1: 2098), and the COICOP has only 157 categories, but is restricted to HH final consumption of goods and services only. The number of categories of the international ISIC is in between that

[5]The paper makes reference to several classifications: i.e. ISIC (International Standard Industrial Classification), CPC (Central Product Classification), HS (Harmonized System), SITC (Standard International Trade Classification), EBOPS (Extended Balance of Payments of Services classification), BEC (classification by Broad Economic Categories), COICOP (Classification of Individual Consumption by Purpose), COFOG (Classification of Functions of Government), COPNI (Classification of the Purposes of Non-Profit Institutions Serving Households), and COPP (Classification of the Outlays of Producers According to Purpose). All classifications and correspondences can be accessed through the UNSD WEB site (http://unstats.un.org/unsd/class/default.htm). The Extended Balance of Payments of Services (EBOPS) classification and its correspondence with the CPC is published in the Manual on Statistics of International Trade in Services, published by UN, EU, IMF, OECD, UNCTAD, WTO, 2002, UN doc nr. ST/ESA/STAT/SER.M/86.

[6]The symbols for correspondences used in the text are 1:1 referring to ONE-ONE correspondences, 1:M for ONE-MANY correspondences, M:1 for MANY-ONE correspondences, and N:M for MANY-MANY correspondences.

[7]This paper will focus on the use of the HS in classifying foreign trade data, even though some countries still might use the SITC. The HS is more detailed than the SITC; only M:1 relations characterize the correspondence between the two.

of the CPC and COICOP, i.e. 292 for ISIC version 3 and 298 for ISIC version 3.1. The EBOPS breakdown of exports and imports of services includes 51 detailed categories.

- The HS-CPC correspondence is mainly an M : 1 correspondence, with many categories of the HS corresponding to fewer categories of goods in the CPC (5224–1151). The remaining categories of the CPC are services that are not linked to the HS, but to the EBOPS. The EBOPS-CPC correspondence of services (51–543) has mainly 1 : M relations, but there are several instances (education, health, insurance, sea transport services), where there is an M : 1 EBOPS-CPC correspondence. It should also be noted that the number of CPC service categories corresponding to the EBOPS classification are fewer than the number of services categories in the CPC (950), as not all CPC service categories are considered relevant for exports or imports.
- The CPC-ISIC correspondence, which is based on CPC version 1.0 and ISIC version 3, is also mainly M : 1 (2037–290), as there are much fewer ISIC than CPC categories at the most detailed level of each classification.[8]
- The CPC-COICOP correspondence includes three types of categories: i.e. (i) categories of COICOP that are linked to the CPC mainly in a 1 : M relation (116–819); (ii) goods and services not destined for final consumption (1281 CPC categories) that are only included in the CPC and not in the COICOP; and (iii) COICOP categories (41) of individual (as distinct from collective) consumption of government and NPIs that are not linked to the CPC.
- The BEC correspondence table with the HS not only makes a distinction within the HS between intermediate consumption, HH final consumption, and gross fixed capital formation, but also includes further detail within each of these three SNA categories, which permits a more precise correspondence with the HS and CPC. At the most detailed level it includes 19 categories, which refer among others to "food and beverages, distinguishing between primary, processed (products) and (products) for industrial and household use," "consumer goods n.e.c., durable, semi-durable and non-durable," "industrial supplies, primary and processed," "transport equipment, passenger motor cars, other, parts and accessories," etc. The HS has mostly M : 1 correspondences with the three BEC categories: of the HS categories 2850 correspond to intermediate consumption, 1246 to HH final consumption, and 954 to capital formation. A fewer number of HS categories, however, are linked to more than one category of the BEC (N : M relation), including 16 mixed categories of intermediate and HH final consumption, such as fuels, 136 mixed categories of HH final consumption and gross fixed capital formation, such as cars and other transport equipment, and 19 mixed categories of all three expenditure categories, covering non-specified goods (n.e.c.). At present there is no correspondence defined between the EBOPS classification and the BEC.

[8] Of the 292 categories of ISIC 3, only 290 have CPC 1.0 correspondences and these account for 2037 out of the 2098 categories of the ISIC.

2.2. *Adapting International Classifications and Correspondences to a National "System of Classifications"*

When further detailing the basic scheme, as presented in Table 1, one might incorporate all international classifications introduced in the previous section, as well as all cross-tabulations between those classifications. In practice, it is not possible to include all cross-classifications, because of data restrictions and lack of analytical interest. One reason, which relates both to compilation and analysis, is that multiple dimension matrices are difficult to handle because of their size. For instance, if output were cross-classified by ISIC, CPC and BEC, using, say, 20 categories of ISIC, 50 categories of CPC, and the three categories of the BEC, this would result in a multidimensional matrix with 3000 (50 × 20 × 3) elements. Similarly, the intermediate consumption matrix of the SUT would have 1000 categories, if only ISIC and CPC are used. The large number of information items resulting from this would be difficult to handle in analysis, and also surveys and other data sources may not support the large data detail required by such comprehensive output and input matrices. Furthermore, many elements may be either zero or very small and not significant enough in value for use in analysis.[9]

Simplification of the SUT matrices of cross-classified data should therefore be a principal objective in compilation and analysis. Simplification, however, should not be the only objective. It should go hand in hand with developing the SUT framework into an optimal instrument in the compilation and analysis of data. There are several ways of accomplishing this in specific country settings, and classifications and their correspondences play a major role in it. In the sections below are discussed criteria for the coordination of classifications; they may be taken into account when adapting international classifications and correspondences in a coordinated manner and arrive at a "system of classifications" for use in countries. In Section 2.3 the criteria are brought together and reflected in what will be called a simplified SUT framework.

The criteria are aimed at coordinating groups of classifications: ISIC-Sector (Section 2.2.1), CPC-HS-COICOP (2.2.2), ISIC-CPC (2.2.3), BEC-HS-CPC (2.2.4). If, on the basis of one of the criteria, one or more classifications are adapted to local circumstances, adaptations of other classifications may be needed as well, given the interlocking nature of the various classification criteria.

2.2.1. ISIC-Sector Correspondences

Two features of the organization of production in the country are brought out through the classification of data at the meso level of the SUT. The first one is the institutional integration of production units (establishments) and the second is the technical integration of production processes. The institutional classification—reflected in NFC, FC, GOV, etc. distinctions of sectors—was already

[9]These types of matrices are referred in the literature as "sparse" matrices (see Young, 1971). It is argued in that literature that calculations based on inverses and other manipulations of large sparse matrices lead to inaccurate results, due to the incorporations of the large number of zeros and other small values, and that use of approximations of the matrices, which avoid the large number of zero elements, may actually lead to improved results. National accountants may take these considerations into account when designing and compiling the SUT.

embedded in Table 1. The technical integration of production processes is represented at the meso level through the traditional use of the ISIC or industry breakdown of GDP and other production related variables. The correspondence between the two classifications of production related variables shows how the institutional organization of production relates to its technical organization. This correspondence was incorporated in the 1993 SNA in a separate table, the Cross Classification by Industries and Institutional Sectors of Production Account Items (CCIS), and thus given for the first time an analytical meaning and quantification in national accounting. Thus can be quantified the percent of production in manufacturing that takes place in small (HH) establishments, large corporations (NFC) and in the government (GOV) sector. Also within production by NFCs can be distinguished the part that takes place in public, foreign controlled and other private NFCs.

The objective of coordinating the ISIC and sector classifications effectively in an analytical SUT framework, in which production is analyzed in a particular country setting, is to aggregate or disaggregate sectors and ISIC categories in such a manner that the link between the institutional and technical organization of production at the micro level of the establishment is well reflected at the meso level of the SUT. Thus, the detail of sector and ISIC classifications in the CCIS at the meso level should bring out that selected establishments are predominantly organized institutionally, for instance as small (HH) or large (NFC or FC), form part of the government sector (GOV) or a KEY sector of the economy. If the characteristics of these establishments are not well represented in the ISIC and/or sector breakdown, the particular organization of production will not be brought out at the meso level of the data.

A simple and transparent CCIS matrix framework should include ideally only 1 : 1 and M : 1 correspondences between ISIC and sector categories, i.e. each ISIC category corresponds to only one sector category. In this case, the sector classification of data can be derived from the ISIC detail of the SUT framework by simple aggregation. This supports the integration of the SUT data with the other segment of the SNA, i.e. the IEA. From an analytic point of view, it has the further advantage that clear links can be established between the technical analysis of the SUT and more institutional analyses of the economy that do not only include analysis of production, but also fiscal, monetary and financial analyses.

In order to arrive at the ideal case as closely as possible, both the ISIC and sector breakdowns would need to be adapted to local circumstances. Thus, if for instance there were two types of agriculture in the country, i.e. plantation and subsistence agriculture, this should be brought out in a national adaptation of the ISIC breakdown, so that the first one would be allocated to the NFC sector and the second one to the HH sector. Similarly, the introduction of a KEY sector in the sector breakdowns would make it possible in an oil producing county to distinguish between the key sector of oil production and production units producing other minerals, which are NFCs, but do not belong to the key sector. Different sectors also could be distinguished on the basis of compilation criteria. Thus, a distinction may be made between industries on which direct survey information is available and that may be allocated to what may be called a “registered” NFC sector, and industries for which estimates are made on the basis of assumptions about employment-output ratios and that may be allocated to a "non-registered" NFC sector.

2.2.2. Use of the CPC as the Core Product Classification

The CPC is the main product classification of the SUT. To establish it as the core classification requires that it be possible either to convert the other product classifications to the CPC through an M : 1 correspondence or to convert the CPC into a more summary classification through an 1 : M correspondence between that classification and the CPC. The first type is represented by the M : 1 correspondence between the HS classification of exports and imports and the CPC used to classify output, and the second type by the 1 : M correspondence between the COICOP classification of HH final consumption data obtained from HH surveys and the CPC classification of output.

If countries use the detail of the international HS and CPC classifications, the above condition should not present any difficulty in the actual practice of countries, as the present HS-CPC correspondence only includes M : 1 relations. There would also not be any difficulty, if the actual CPC detail of output at the country level were less than the number of categories included in the international version of CPC; in that case, international HS-CPC correspondences may simply be further aggregated. A problem may be encountered, however, if a country has less HS data detail than the 5219 HS categories that have correspondences with CPC 1.0.[10] As this may result in 1 : M and N : M HS-CPC correspondences, the HS detail of exports and imports and/or CPC detail of output and the correspondences between them may need to be adjusted.

This may be done after comparing the actual values of HS detail of imports and exports with the corresponding values of CPC detail of output. Based on this comparison, quantitatively small categories in either one or both of the two classifications may be aggregated, or alternatively larger categories disaggregated, so that actual 1 : M and N : M HS-CPC correspondences are converted into a 1 : 1 correspondence between the two classifications. If the CPC detail is reduced in line with HS export and import data availability, one should take into account the impact that such reduction has on the CPC-COICOP correspondence dealt with in the next paragraph, and on the CPC-ISIC correspondence discussed in the next section.

The conditions set out above are also satisfied in the present international COICOP-CPC correspondence, which links output to HH final consumption. The international COICOP has much fewer categories (116) than the final consumption categories that can be identified in the international CPC (819), there is an M : 1 correspondence between the CPC and COICOP. If the detail of the CPC is adjusted at the national level, because some product categories are eliminated as they are not relevant nationally, detail has been added for national purposes, or CPC categories have been aggregated because data detail is not available—the

[10]This may occur when countries are still using the SITC or a previous version of the HS, i.e. the BTN (Brussel's Trade Nomenclature), to classify foreign trade. Both classifications have less detail than the present HS. The SITC has 3122 categories at the most detailed level, compared with the 5224 categories of the HS.

CPC-COICOP correspondence should be also adjusted so that the M : 1 correspondence is maintained.

Another consistency requirement that should be met relates to the BEC-CPC correspondence. If the BEC is extended to goods and services of the CPC, as is suggested in Section 2.2.4 below, it would identify within the CPC all goods and services that correspond to HH final consumption. This scope, of course, would need to be made compatible with the scope of the 819 CPC categories of HH final consumption that are identified through the CPC-COICOP correspondence.

2.2.3. Coordinate ISIC-CPC Correspondences

With regard to the technical integration of production data, the objective is to coordinate ISIC and CPC detail in such a manner that the meso SUT reflects as well as possible the production structure at the micro level. This refers in particular to the distinction between principal and secondary products, which can be only made well at the micro level of the establishment. If aggregations of ISIC and CPC categories are not well coordinated, at the meso level a distinction between characteristic and non-characteristic products results, which does not match the micro distinction between principal and secondary products. In particular non-characteristic products at the meso level may identify not only secondary products, but also other non-characteristic products that are only included as a result of the aggregation of ISIC and CPC categories.

The distinction between principal and secondary products can be made both in a square and also in a rectangular ISIC-CPC output matrix. In a square output matrix the diagonal elements refer to principal products, and the non-diagonal elements should only include secondary products that are identified as such at the micro level. In a rectangular matrix the secondary products identified at the micro level should show up at the meso level as the entries with minor values in each industry column. A more accurate distinction between principal and secondary products can be made at the macro level by introducing more detail in the CPC breakdown and thus making the output matrix more rectangular.

The percentage shares of non-characteristic products in total output at four different levels of aggregation of ISIC and CPC categories may illustrate the impact of ISIC and CPC aggregation.[11] At the most detailed level of both (13 ISIC and 23 CPC categories), the share of non-characteristic products is 9.3 percent. If one assumes that the output table at this level of aggregation is a fair representation of the distinction of principal and secondary products at the micro level, the share of non-characteristic products only refers to secondary products, and characteristic products are principal products. When the number of CPC categories is aggregated to 10 only, while leaving the number of ISIC categories (13) unchanged, the share of non-characteristic products decreases to 6.2 percent, as some secondary products that were identified at the previous level of aggregation, are now included with the value of characteristic products. On the other hand, if the

[11]The example is based on the illustrative output data presented in table 15.1 of Chapter XV of the 1993 SNA.

number of ISIC categories is reduced from 13 to 8 (leaving the number of CPC categories at 23), the share of non-characteristic products increases dramatically to 44.1 percent. This is because products that were identified as principal products at the previous level of aggregation, are now included with non-characteristic products. If both ISIC and CPC are aggregated to fewer categories (8 ISIC and 10 CPC categories), the share of non-characteristic products is somewhere in between that of the two previous values of the shares, i.e. 15.1 percent. In all instances except the first one, the share of non-characteristic products differs from the share of secondary products, which was assumed to be 9.3 percent as measured in the most detailed output table. Whether the assumption is correct that the distinction between principal and secondary products is well reflected at the meso level of the first ISIC/CPC aggregation, of course, can only be verified after a detailed quantitative investigation of the scope of secondary products at the micro level.

2.2.4. Use the BEC to Trace Imports and Outputs to Uses

Once the CPC is recognized as the core classification of the SUT, further refinements could be introduced, when applying the international classification in countries the refinement may include:

- Distinctions, where possible, within the CPC categories between products that are imported and those that are produced in the country.
- Redefinition of the present BEC-SITC link to a BEC-CPC link and thus cover services as well as goods, and apply the BEC not only to imports, but also to output.
- Limitation of the number of CPC categories that corresponds to more than one BEC category.
- Extension of the BEC from the traditional distinction between intermediate consumption, HH final consumption and capital formation, to include—at the country level—also links with exports, GOV and NPI consumption.

If these refinements are introduced to a maximum extent, most products would have one origin—imports or output—and only one of the destinations mentioned. These constitute the five attributes (p, m, ci, c, k and x) that are indicated in the scheme in Table 2. To these are added the distinction between goods and services (g,s), as that is a distinction which is made between two foreign trade classifications, i.e. HS and EBOPS. These seven attributes are used in the center of the table to identify the attributes for reclassifying the data of statistical sources for the collection of data on supply (row) and use (column) of products. The headings of rows and columns indicate respectively the data sources, SNA items and type of classification categories that need to be reclassified in this manner. For instance in the second row on the supply side, foreign trade merchandise data on imports (m) from the customs register that are classified by HS, should be reclassified into three groups, i.e. imported (m) goods (g) destined for intermediate consumption (m/g/ci), capital formation (m/g/k) and HH final consumption (m/g/c). Similarly in the last column on the use side, data on HH final consumption (c) from the HH survey should be reclassified into four groups, i.e. HH final

TABLE 2

ATTRIBUTES OF PRODUCT FLOWS IN THE SUT

	Use →	Data source	Economic survey, GOV registers	Customs register	BOP	Household survey
	supply ↓	SNA item(s)	Intermediate consumption, capital formation	Exports of merchandise (goods)	Exports of services	HH final consumption
Data source	SNA item(s)	Classification	CPC	HS	EBOPS	COICOP
Economic survey, GOV registers	Output	CPC	p/g/ci p/s/ci p/g/k p/s/k	p/g/x	p/s/x	p/g/c p/g/c
Customs register	Imports of merchandise (goods)	HS	m/g/ci m/g/k			m/g/c
BOP	Imports of services	EBOPS	m/s/ci m/s/k			m/s/c
Data attributes	p	Output	c	HH final consumption	g	Goods
	m	Imports	k	Gross fixed capital formation	s	Services
	ci	Intermediate consumption	x	Exports		

consumption goods domestically produced (p/g/c) and imported (m/g/c) and HH final consumption services domestically produced (p/s/c) and imported (m/s/c).

If all classification categories corresponding to the SNA items in the data sources indicated were used on both the supply and use side, it would result in a maximum number of data checks between supply and use, as the same data categories would be available on both sides.

In order to accomplish this, however, existing international classifications and correspondences, when used in countries, would need to be:

- Utilized in the maximum detail feasible. If less detail is used, regrouping of the data from the sources as indicated would be more difficult.
- Amended in line with products that are locally produced or used, and for which the origin and destination can be clearly identified. Thus HS and EBOPS would need to be adapted for local use, so that these classifications only include categories of goods and services that are actually imported or exported by the country, and COICOP would need to be amended for national use so that it includes only products that are locally consumed by HHs.

What is described above may be called an ideal situation, which in practice may be difficult to accomplish, given data restrictions. Already in the present BEC there are several categories of HS, which are either used for HH consumption or capital formation, HH consumption or intermediate consumption, or for either one of the three categories, i.e. HH consumption, intermediate consumption or capital formation. If the detail of the HS as used in a country were less than the detail available in the international HS, the number of product categories with combined distinctions may be increased further. On the other hand, if an effective adaptation of the HS to local circumstances were made, this problem may be mitigated. Similarly, it may difficult on the use side to distinguish between domestically produced products and imported products. But also in this case, the problem may be reduced by amending the classifications on the use side to locally produced or used products, for which it would be easier to determine which ones are imported and which ones are domestically produced.

2.3. *Classification Refinements Incorporated in a Simplified SUT Framework Within the Extended SNA*

The framework design and the classification refinements discussed above have been reflected in Table 3 in an operational format of the SNA-SUT. The SUT presented is comprehensive in the sense that it includes all classifications and cross-classifications that might be used to refine the detail of the SUT. On the other hand, it is referred to as a simplified SUT framework, as it includes various simplifications, particularly in the cross-classification of data. The SUT table has been designed for purposes of compilation and analysis of data, and may thus be used in the actual SNA compilation in countries.[12]

The features of the SUT that derive directly from Table 1 are reflected in the columns of the sectors that are identified in the left hand columns of Table 3, i.e.

[12]The author has developed simplified SNA frameworks for Guyana, Oman, Nicaragua and St. Vincent, along the lines of Table 3.

Table 3 Simplified SUT

	Sector/ISIC	NFC	FC	GOV	HH	NPI	NFC/HH	NFC	NFC/HH	NFC	NFC/HH	HH	HH	NFC	NFC/FC/HH	GOV	NFC/GOV/HH/NPI	Total Economy and ROW
							Agriculture, forestry, fishing (A,B)	Mining and quarrying (C)	Manufacturing (D)	Electricity, gas and water supply (E)	Construction (F)	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods (G)	Hotels and restaurants (H)	Transport, storage, and communication (I)	Financial intermediation, real estate, other business services, private household services (JPK)	Public administration and defence; compulsory social security (L)	Education, health, other community, social and personal services (MNO)	
	CPC corres-pondence						01-04	11-16	21-51	69	53-54	61-62	63	64-68, 84	71-83, 85-87	91, 99	92-98	
	Product destination						IC_{hh}	I	$IC_{hh}KX$	IC_{hh}	IK	IC_{hh}	X	$IC_{hh}X$	IC_{hh}	C_{gov}	$IC_{hh}C_{gov}C_{npi}$	
	Product origin						PM	PM	PM	P	P	PM	P	PM	PM	P	PM	
Total output	Sector/ISIC/CPC	**1,753**	**102**	**440**	**1,269**	**40**	**89**	**35**	**1,680**	**164**	**244**	**123**	**68**	**100**	**478**	**168**	**455**	**3,604**
CPC detail 21-25	CPC								448									448
CPC detail 26-31	CPC								422									422
CPC detail 32-37, 41-42	CPC								796									796
CPC detail 64	CPC													44				44
CPC detail 66	CPC													31				31
CPC detail 72-83, 85-87	CPC																98	98
CPC detail 92-94	CPC																276	276
CPC detail 95-98	CPC																79	79
Imports	CPC						**37**	**60**	**283**			**79**		**20**	**13**		**5**	**497**
Taxes less subsidies on products	CPC						2	0	89	5	17	0	3	5	8	0	4	133
Trade & transport margins	CPC						2	2	74	0	0	-68	0	-10	0	0	0	0
Supply at market prices	CPC						**130**	**97**	**2,126**	**169**	**261**	**134**	**71**	**115**	**499**	**168**	**464**	**4,234**
Intermediate consumption (directly	Sector/ISIC	899	29	252	694	9	47	17	1,000	127	114	64	36	60	132	118	168	1,883
Value added	Sector/ISIC	854	73	188	575	31	42	18	680	37	130	59	32	40	346	50	287	1,721
Employment	ISIC						2,058	292	28,682	3,300	5,024	5,000	2,078	2,032	3,700	8,000	9,203	69,369
Changes in inventories	ISIC						1	-1	15	0	23	0	0	0	0	0	0	38
Gross fixed capital formation (directly estimated by industry)	ISIC						11	6	104	13	9	13	7	39	144	12	18	376
Intermediate consumption	CPC						88	96	992	123	40	61		78	309		95	1,882
HH Final consumption	CPC				1,004		28		567	36		51		14	250		58	1,004
GOV Final consumption	CPC			363				0			0				0	159	204	363
NPI Final consumption	CPC					14	0	0	0		0				0		14	14
Gross fixed capital formation (allocated by product)	CPC							0	161		190						0	351
Changes in inventories	CPC						1	-1	15		23				0		0	38
Exports	CPC								**422**				**67**	**19**				**508**
Use at market prices	CPC						**117**	**95**	**2,157**	**159**	**253**	**112**	**8**	**111**	**559**	**159**	**371**	**4,101**
Supply-use discrepancies	CPC						**13**	**2**	**-31**	**10**	**8**	**22**	**4**	**4**	**38**	**9**	**-5**	**74**

NFC, FC, GOV, HH and NPI. The additional columns on the right hand side refer to the ISIC categories of industries that were discussed in Section 2.2.1, and that traditionally are used for the industry breakdown of GDP and other production related variables. The right hand column of the table is a combination of the columns for the National Economy and ROW in Table 1. Thus it includes all totals, which are aggregate data of resident sectors, and also some ROW data, particularly exports and imports. Most items are specified by sectors and their totals are included in the National Economy & ROW column: i.e. output, intermediate consumption, value added and final consumption by HH, GOV and NPI sectors. Some items are only available in totals and presented only in the National Economy & ROW column, i.e. gross fixed capital formation, changes in inventories, and taxes less subsidies on products. The supply-use identity holds, but only for the items in the National Economy & ROW column and not for the sectors individually.

The CCIS cross-classification of ISIC and sector categories is incorporated by indicating in the first row above the industry groupings, the sectors to which they correspond.[13] For instance, the ISIC category "Agriculture, forestry, fishing (A,B)" corresponds to the NFC and HH sectors, and the ISIC category "Public administration and defence; compulsory social security (L)" is allocated to the GOV sector. The CCIS correspondences only cover the items of output, intermediate consumption and value added, which are classified alternatively by sector on the left hand side and by ISIC categories on the right hand side of the table. The totals of the three categories are the same for both types of classifications (respectively 3,604, 1,883, 1,721). Other items, including employment, changes in inventories, gross fixed capital formation (directly estimated by industry), have been classified by ISIC categories only.

The CPC breakdown has been incorporated in the table in two ways. Firstly, CPC categories have been included as correspondences to ISIC categories at the heading of each ISIC column. Thus, CPC categories 01–04 correspond to ISIC category Agriculture, forestry, fishing (A,B), or CPC 53–54 corresponds to ISIC Construction (F). The ISIC-CPC correspondence is used to incorporate in the columns on the right had side of the table not only the production categories of output, intermediate consumption (directly estimated by industry), value added, employment, changes in inventories (directly estimated by industry) and gross fixed capital formation (directly estimated by industry), classified by ISIC categories, but also the supply and use categories of output, imports, intermediate consumption (allocated by product), HH final consumption, gross fixed capital formation (allocated by product), changes in inventories (allocated by product), and exports, which are classified by CPC categories. The right hand side of the table thus integrates the ISIC columns and CPC rows of the traditional SUT, using the assumption that there is a correspondence between ISIC and CPC categories. The distinction between the ISIC and CPC categories in the table has been made by shading the rows corresponding to the CPC product categories. This needs two qualifications. The first one is that non-shaded output is classified by both ISIC and CPC, and the other one is that there are three items of intermediate con-

[13]The ISIC-sector correspondences are based on those included in table 15.3 of the 1993 SNA.

sumption, gross fixed capital formation and changes in inventories, i.e. one for each directly estimated by industry (ISIC), and another one allocated by product (CPC).

In the table, as in practice, the assumed ISIC-CPC correspondence only holds approximately, and this is reflected in the statistical discrepancy at the bottom of the table between supply and use. In order to mitigate the impact of this assumption on the accuracy of the SUT, further refinements have been introduced in a second appearance of CPC categories in the rows below total output, reflecting the classification coordination criteria developed in Section 2.2.3. In each industry column have been presented among others the secondary product categories that are included in the output of one industry, but characteristically belong to another industry. An example is in the last column of "education, etc. (ISIC:MNO/CPC: 92–98)," where a figure of 98 appears as secondary output corresponding to CPC 72–83, 85–87, which characteristically corresponds to ISIC category "Financial intermediation, real estate, other business services, private household services (JPK)." With this CPC refinement in the rows, supply of CPC category 72–83, 85–87 has been augmented with secondary output produced elsewhere (597 = 478 + 98 + 13 + 8), and supply of CPC category 92–98 was correspondingly reduced (388 = 455 − 98 + 5 + 4). By introducing these refinements in the CPC rows related to secondary output the statistical discrepancy at the bottom of the table is reduced. The advantage of this approach is that a comprehensive output table is not needed, and only a selection of quantitatively important secondary outputs is introduced as rows below total output. This reduces the amount of work of national accountants on quantitatively unimportant secondary output categories, improves the reliability of the quantitatively important ones and avoids the large number of zero (0) and very small data elements presently included in large output matrices.

The CPC rows have not only been used to identify secondary output, but also to identify products that would help to trace better the correspondence between supply and use of products as discussed in Section 2.2.4. Thus, for instance in the column of "Manufacturing (D)" have been distinguished three product categories: 21–25 (Food and tobacco), 26–31 (Textile, leather and wood products), 32–37, 41–42 (Low-tech manufacturing products for industrial use).[14] The first group of products is predominantly destined for HH final consumption, the second group for exports and the third group for intermediate consumption, according to a BEC adapted to the country in question. By making this additional CPC distinction it is thus possible to allocate from the supply side output to its corresponding use category and compare those estimates with the direct estimates of each destination category. Further CPC detail could also be incorporated in the case of imports, when the total value of imports recorded in each column does not correspond in its entirety to the CPC category indicated in the heading of the column. Additional ISIC detail could furthermore be incorporated in additional rows below "Intermediate consumption (allocated by product)," if for selected ISIC categories more CPC detail on intermediate inputs is available from surveys and other data sources.

[14]The category "Low-tech manufacturing products for industrial use" referring to categories 41–42 is not an aggregate CPC category, but has been created solely for the purpose of this paper.

3. Conclusion

The simplified SUT framework including the use of classifications and cross-classifications to introduce further detail is flexible enough to incorporate various analytical options. Thus, one should keep in mind that not all details suggested above are to be implemented in practice when designing the SUT framework. This conclusion applies not only to the limited segment of the SUT, but also to the larger SNA framework and to analytical frameworks in general. Additional details require additional data, and this may sometimes be too costly and/or too time consuming, as compared to the analytical benefits obtained. Therefore the actual format of a framework should be decided in close cooperation between national accountants and policy analysts, who can only determine together what would be the most optimal scheme from a statistical and analytical point of view. Thus, in the case of the SUT, they may need to decide whether the ISIC breakdown be used solely without CPC detail, in which case it should be accepted that only limited data checks and analytical indicator ratios are available. Or alternatively, more detailed CPC breakdowns may be introduced, which increases the data requirements of the framework, but also increases the number of data checks and the analytical potential of the framework. Incorporating cross-classifications between ISIC and sectors would further enhance the analytical potential of the framework, as it would allow joining GDP growth analysis based on ISIC and CPC data of the SUT with fiscal, monetary and financial analyses based on the IEA.

References

Commission of the European Communities–Eurostat, International Monetary Fund, Organization for Economic Cooperation and Development, United Nations, World Bank, *System of National Accounts, 1993* (United Nations publication sales no. E.94.XVII.4).

FAO, *System of Economic Accounts for Food and Agriculture*, 1996 (ISBN 9251038074, Job Number W0010/E).

Hoffmann, E., "Developing Labour Account Estimates: Issues and Approaches," in *Household Accounting: Experiences in the Use of Concepts and their Compilation. Vol. 2: Household Satellite Extensions, Handbook of National Accounting,* United Nations, 2000.

ILO, *Compendium of Official Statistics on Employment in the Informal Sector*, by Ralf Hussmanns and Brigitte du Jeu, STAT Working Paper No. 1, Geneva, 2001.

IMF Statistics Department, *Balance of Payments Manual*, Washington, D. C., September 1993.

———, *Monetary and Financial Statistics Manual*, Washington, D. C., September 2000.

———, *Government Finance Statistics Manual*, Washington, D. C., 2001.

Korea (Republic of) National Statistical Office (Kyung Hee Kim, Eun Pyo Hong, Eun Soo Lee), UNSD (Jan W. van Tongeren and Berndt Becker), UNDP, *Human Resource Accounts for Korea: A Study of Education and Employment*, Seoul, October 2000.

Magnus, Jan, CentER Tilburg University, Jan W. van Tongeren, UNSD, Aart F. de Vos, Free University, Amsterdam, National Accounts Estimation Using Indicator Ratios, *Review of Income and Wealth*, 6(3), September 2000.

Ruggles, Richard and Nancy, D. Ruggles, "The Role of Micro Data in the National Economic and Social Accounts," *Review of Income and Wealth*, 21(2), June 1975.

Ruggles, Nancy, D. and Richard Ruggles, "The Treatment of Pensions and Insurance in the National Accounts," *Journal of the International Association for Research in Income and Wealth*, 29(4), December 1983.

———, "The Integration of Macro and Micro Data for the Household Sector," *Review of Income and Wealth*, 32(3), September 1986.

———, "Household and Enterprise Saving and Capital Formation in the United States, A Market Transactions View," *Review of Income and Wealth*, 38(2), June 1992.

Thorbecke, Erik, Cornell University, “The Use of Social Accounting Matrices in Modeling,” paper presented at the International Association for Research in Income and Wealth 26th General Conference, Cracow, Poland, 27 August to 2 September 2000.

United Nations, *Manual X, Indirect Techniques for Demographic Estimation*, 1983 (sales no. E81.XIII.7).

———, *Integrated Environmental and Economic Accounting,* 2004.

Van Tongeren, Jan, “Development of an Algorithm for the Compilation of National Accounts and Related Systems of Statistics,” *Review of Income and Wealth*, 22, March 1986.

Young, David M., *Iterative Solutions of Large Linear Systems*, Academic Press, London and New York, 1971.

Chapter 3

Integrated environmental and economic accounting: Framework for a SNA satellite system

Review of Income and Wealth
Series 37, No. 2, June 1991

INTEGRATED ENVIRONMENTAL AND ECONOMIC ACCOUNTING: FRAMEWORK FOR A SNA SATELLITE SYSTEM

BY PETER BARTELMUS
United Nations Statistical Office

CARSTEN STAHMER
Federal Statistical Office, Germany

AND

JAN VAN TONGEREN
United Nations Statistical Office

National accounts have provided the most widely used indicators for the assessment of economic performance, trends of economic growth and of the economic counterpart of social welfare. However, two major drawbacks of national accounting have raised doubts about the usefulness of national accounts data for the measurement of long-term sustainable economic growth and socio-economic development. These drawbacks are the neglect of (a) scarcities of natural resources which threaten the sustained productivity of the economy and (b) the degradation of environmental quality from pollution and its effects on human health and welfare. In the present paper, the authors attempt to reflect environmental concerns in an accounting framework which maintains as far as possible SNA concepts and principles. To this end, the accounting framework is used to develop a "SNA Satellite System for Integrated Environmental and Economic Accounting" (SEEA). Environmental costs of economic activities, natural asset accounts and expenditures for environmental protection and enhancement, are presented in flow accounts and balance sheets in a consistent manner, i.e. maintaining the accounting identities of SNA. Such accounting permits the definition and compilation of modified indicators of income and expenditure, product, capital and value added, allowing for the depletion of natural resources, the degradation of environmental quality and social response to these effects. A desk study of a selected country is used to clarify the proposed approaches, to demonstrate their application in future country studies and to illustrate the quantitative effects of the use of modified concepts on the results of analysis.

1. INTRODUCTION

The discussion of environmentally sound and sustainable socio-economic development has received increased attention by the international community, stimulated in particular by the report of the World Commission on Environment and Development (1987). At its forty-second session, the General Assembly welcomed the Commission's report (resolution 42/187) and adopted an "Environmental Perspective to the Year 2000 and Beyond" which proclaimed "as the overall aspirational goal for the world community the achievement of sustainable development on the basis of prudent management of available global resources and environmental capacities" (resolution 42/186). Environmentally sound and

Note: The authors thank Hubert Donnevert and Stefan Schweinfest for their assistance on the present version. The views expressed by the authors are their own and not necessarily those of their respective institutions.

sustainable development will also provide the basic theme for the planned United Nations Conference on Environment and Development in 1992.

The need for clarifying this new development concept and for developing methodologies for its assessment and implementation has been recurrently stressed in international conferences, seminars and workshops. Joint workshops, organized by UNEP and the World Bank, examined the feasibility of physical and monetary accounting in the areas of natural resources and the environment and developed alternative macro-indicators of ecologically adjusted and sustainable income and product (Ahmad, El Serafy and Lutz, 1989). A consensus emerged in the workshops that enough progress had been achieved to link environmental accounting to the standard System of National Accounts, the SNA (United Nations, 1968), and to include certain aspects of environmental accounting in the ongoing revision of SNA.

National accountants and environmentalists reviewed a first draft of the present paper in a UNEP/World Bank-sponsored expert meeting (Paris, November 21-22, 1988). The experts at the meeting endorsed the idea of developing a satellite system of environmental accounts and discussed a variety of methodological and procedural questions. These questions should be resolved before preparing an internationally recommended manual of environmental accounting. The experts also requested that the revised SNA should elaborate on the approaches to incorporating environmental concerns in national accounts.

The immediate objective of the present framework is to serve as the basis for the preparation of a "SNA Handbook on Integrated Environmental and Economic Accounting" to be issued within the United Nations series of national accounting handbooks. The framework should also facilitate the consideration of environmental accounting in the revised SNA, possibly as part of a more general treatment of the concept of satellite accounts and with appropriate cross-referencing to the Handbook. The draft methodologies have been tested in pilot country studies, and will be distributed widely for comments and contributions.

The framework discussed in this paper is the basic structure for a "Satellite System for Integrated Environmental and Economic Accounting" (SEEA). It is presented in tabular form with an illustrative set of data and is described in some detail in the text. In section 2 the main objectives of environmental accounting as well as the general structure of the SEEA are described. Section 3 contains a description of the supply side of goods and services, focusing on environmental protection services and the supply of natural growth products. The accounting for the costs of environmental depletion and degradation, resulting from production and consumption, is the main issue of section 4. In this section, the authors also explain how these costs affect value added and final demand. One basic indicator, the Environmentally Adjusted Net Domestic Product or "Eco Domestic Product" (EDP) is presented in this context. In section 5 the flow accounts of sections 3 and 4 are complemented by the presentation of stock assets of tangible wealth that include natural assets and changes therein. In section 6 the possible extensions of the flow accounts to obtain welfare-oriented macro-indicators are discussed. Finally, some comparative analyses of the conventional and environmentally modified concepts are presented in section 7.

2. General Features of a Satellite System for Integrated Environmental and Economic Accounting (SEEA)

(a) *Objectives of Integrated Environmental and Economic Accounting*

The focus of traditional systems of national accounts on market and some related non-market transactions (except for imputations for "directly competitive" non-market production of goods and services) has effectively excluded the accounting for changes in the quality of the natural environment and the depletion of natural resources. These effects have been considered to be particularly relevant for the assessment of long-term sustained growth and development and of increases in "social welfare." The overall objective of environmental accounting is thus to measure more accurately the structure, level and trends of socio-economic performance for purposes of environmentally sound and sustainable development planning and policies. The attainment of this objective would facilitate both the systematic compilation and analysis of environmental and related socio-economic data and the formulation of alternative standard macro-economic variables for the analysis of environmental-economic interrelationships.

The current revision of the SNA (United Nations, 1990) presents a unique opportunity to examine how the various concepts, definitions, classifications and tabulations of environmental and natural resource accounting can be linked to or incorporated in the SEEA. It may appear premature, however, to radically change a well-established system of economic accounts that serves many different short-, medium-, and long-term socio-economic analyses. Further elaboration of the standards of environmental and natural resource accounting in a *SNA satellite system* of environmental accounts has therefore been proposed (Bartelmus, 1987). A similar view was expressed by the experts working on the current revision of the SNA (Lutz and El Serafy, 1989).

Satellite systems of national accounts generally stress the need to expand the analytic capacity of national accounting for selected areas of social concern in a flexible manner, without overburdening or disrupting the "core" system (Lemaire, 1987; Teillet, 1988; Schäfer and Stahmer, 1990). Typically, satellite accounts allow for the:

—provision of additional information on particular social concerns of a functional or cross-sectoral nature,

—linkage of physical data sources and analysis to the monetary accounting system,

—extended coverage of costs and benefits of human activities, and

—further analysis of data by means of relevant indicators and aggregates.

Accordingly, the following specific objectives can be formulated for the planned SEEA:

(i) Segregation and Elaboration of All Environment-Related Flows and Stocks of Assets of Traditional Accounts

Satellite accounts, in the narrow sense of detailed accounting for expenditures and revenues in major areas of social concern, were pioneered by France (Institute National, 1986a). There is now an increased interest in segregating all flows and stocks of assets in national accounts related to environmental issues and, in particular, in estimating the total expenditure for the protection or enhancement of the different fields of environment. One objective of this segregation is the identification of the increasing part of the Gross Domestic Product (GDP) which reflects the costs necessary to compensate the negative impacts of economic growth ("defensive expenditures") rather than increases in "true" (welfare-relevant) income (Hueting, Leipert, 1987; Leipert, 1989; and Olson, 1977).

(ii) Linkage of Physical Resource Accounting with Monetary Environmental Accounting and Balance Sheets

Physical resource accounts aim at covering comprehensively the total stock or reserves of natural resources and changes therein, even if these resources are not (yet) affected by the economic system.[1] The proposed accounting for these resources is considered the "hinge" by which comprehensive physical resource accounts could be linked to the monetary balance sheet and flow accounts. Another important method for analyzing the environmental-economic interrelationship in physical terms is the development of material/energy balances (Ayres, Kneese, 1969; Ayres, 1978; United Nations, 1976). This approach allows in particular the linkage of input-output tables with data on natural resource inputs, the description of the transformation of natural resources in the production process and the assessment of the generation of residuals of the economic activities (Isard, 1969; Leontief, 1973). Systems of environmental statistics such as those proposed by the United Nations (in preparation) should facilitate achieving compatibility between physical and monetary accounts by specifying those parameters that could be valued in monetary terms to obtain the figures required in environmental accounts. Non-monetary data in physical accounts are considered to be an integral part of the SEEA and will be fully elaborated in the Handbook on Integrated Environmental and Economic Accounting. However, the present framework will concentrate on the monetary stocks and flows of an environmental accounting system.

(iii) Assessment of Environmental Costs and Benefits

In contrast to the above-mentioned "narrow" satellite accounts, a broader framework for satellite accounting, covering additional "external" environmental costs and benefits, is proposed here. Taking the current state of knowledge and data availability into account, this framework focuses on expanding and complementing the SNA, with regard to two major issues, namely

—the use (depletion) of natural resources in production and final demand and
—the changes in environmental quality resulting from pollution and other impacts of production, consumption and natural events on one hand and environmental protection and enhancement on the other.

Possibilities of extending the framework for the analysis of environmental welfare effects, i.e. the "damage costs" of human health impairment, recreation and other aesthetic or ethical values, are also indicated.

[1]See e.g. the Norwegian approach to natural resource accounting (Alfsen, Bye and Lorentsen, 1987) or the more complex (including interactions in the biophysical environment) French "natural patrimony" accounts (Institute National, 1986).

(iv) Accounting for the Maintenance of Tangible Wealth

The recent discussion of the new paradigm of sustainable development stressed the need to fully account for the use of both man-made and "natural" capital in order to alert to possible non-sustainable growth and development scenarios. The proposed framework aims at extending the concept of capital to cover not only man-made capital, but also natural capital. Accordingly, SEEA will include additional costs for the depletion and degradation of these natural assets. It will also extend the concept of capital formation to capital accumulation which reflects additionally the deterioration of natural capital as a result of economic uses.

(v) Elaboration and Measurement of Indicators of Environmentally Adjusted Income and Product

The consideration of the depletion of natural resources and changes in environmental quality permits the calculation of modified macro-economic aggregates, notably the Environmentally Adjusted Net Domestic Product, short: Eco Domestic Product (EDP).

All these objectives can only be realized step by step. Initial emphasis in practical work should be on the improvement of physical environmental data and on linking them with national accounts as a prerequisite for the valuation of environmental effects.

(b) *Scope and Structure of the SEEA*

The proposed SEEA follows as far as possible the principles and rules established in the SNA (United Nations 1968, 1977, 1990). It is based on SNA's production boundary, follows its analysis of costs and outputs and incorporates the same accounting identities between supply and use of products and between value added and final demand. Information needed for environmental analysis is presented separately. In this manner, original (unadjusted) SNA data can be directly compared with environmentally adjusted statistics and indicators, facilitating the linkage with the central framework of the SNA. Such compliance and linkage with SNA aims at better integration of environmental variables into established economic analysis.

The very nature of a framework allows only the most important concepts and accounting procedures to be highlighted. Definitions, classifications, valuation principles, data sources and processing will be further elaborated on in the Handbook on Integrated Environmental and Economic Accounting. The Handbook will benefit from the experience gained in country studies and existing expertise at the national and international levels.

The present framework seeks to be flexible regarding alternative approaches to integrated environmental-economic accounting and analysis. The interrelationship between the environment and the economy is described as complete as possible. However, in line with the production boundary of SNA, phenomena that take place wholly within the environment, i.e. outside the economic system, are excluded. Such phenomena are probably better accounted for by the use of complementary biophysical resource accounts and systems of environment statistics and monitoring. Also, welfare effects from environmental quality degradation that affect "human capital," i.e. human health and welfare, are not accounted for in the present framework. However, as shown below (see section 6), a "window" to the analysis of environmental damage related to human welfare has been opened, facilitating further extension or alteration of the framework for such analysis.

The main emphasis of the proposed scheme is on the implications of the environment for production, value added, final and intermediate demand and tangible wealth. Therefore, the framework does not present complete accounts for all institutional sectors. Transactions related to income distribution and those concerning intangible assets, including exploitation rights, and also financial assets are excluded. A complete analysis of the interrelationships between the economy and the environment will call for an extended system of all institutional accounts, which shows not only the flows of goods and services, but also of income and finance.

In Table 1 the general structure of the system which consists of three basic components is illustrated. In Tables 1.1 and 1.2 the supply and use of goods and services is shown. The asset accounts with opening and closing assets and the items linking them are shown in Table 1.3. Tables 1.2 and 1.3 are connected via the accounts of capital accumulation. The component tables are further elaborated on in Tables 2, 3 and 5 as explained in sections 3 to 5.

The supply Table 1.1 contains an additional row which shows the involuntary "imports" of residuals (wastes etc.) of foreign economic activities which were transported to the domestic economy (−1.6). The use/value added Table 1.2 is extended by row as well as by column. In the table, we show not only the traditional GDP and NDP, but also further corrections due to the use of natural assets (depletion of natural resources, degradation of natural assets by residuals, agricultural and recreational use etc.). This use is valued with the costs which would have been necessary to keep the natural capital intact (ecological valuation; see below section 4c for an alternative approach in the case of "exhaustible" resources). These costs are interpreted as the decrease in value of the natural assets comparable to the consumption of man-made fixed assets. The deterioration of the natural assets could be caused by current production activities (59.8), consumption activities (household consumption 17.1) or by (scraps of) produced assets (5.1). The restoration activities of the government diminish the impacts of the economic activities on the natural assets (−5.0). The use of natural assets could affect the domestic nature (loss of ecological functions of the produced biological assets −0.9, natural non-produced assets −73.0) or—as far as the generation of residuals is concerned—could lead to transportation to the rest of the world (exports: −4.7). The value of the deterioration of the domestic as well as foreign natural assets caused by domestic sources (59.8 + 22.2 = 82.0) is used for estimating the environmentally adjusted Net Domestic Product (NDP), called Eco Domestic Product (EDP) (185.1) (see section 4c below).

The asset accounts (Table 1.3) show the produced assets (including cultivated biological assets) and the non-produced assets which contain only natural assets (wild biota, land, subsoil assets, water and air). Market valuation is applied except for the depletion and degradation values of natural assets shown in the use/value added table (Table 1.2). These volume changes are valued with the (hypothetical) costs for maintaining them on the same overall quantity and quality level during the reporting period. The question of how such values could be integrated into the asset balances containing mainly market values is discussed in section 5.

3. Supply of Goods and Services

The supply table (Table 2) includes two elements: gross output, resulting from domestic production, and imports. Gross output is cross-classified by industries and type of product (good or service). Imports are classified by the same type of product as domestic gross output, so that the two elements of supply can be added together to obtain total supply by product. Furthermore, the involuntary "imports" of residuals of foreign economic activities are shown. This item could contain e.g. the unaccepted dumping of foreign wastes in national territories.

In Table 2 we show a breakdown of domestic production activities by environmental protection activities and other industries. The fully elaborated system will display a further breakdown by industries according to the International Standard Industrial Classification of all Economic Activities (ISIC) (United Nations, 1990a).

A major modification of the SNA is the separate identification of *environmental protection services* from other production activities for all industries. The separation is to facilitate the assessment of the importance of environmental activities in gross output, employment, other production costs, and in capital consumption. Environmental protection services comprise in principle all activities to maintain and enhance the quality of the natural assets. This could be achieved by avoiding environmental impacts of the economic activities (e.g. by using integrated or end-of-pipe technologies) or by restoring the natural environment already degraded or depleted. Environmental protection activities can be produced for third parties (external use) as main or secondary production activities of the establishments (36.2) or they can be used internally. The internal provision of environmental protection services is considered to be an "ancillary" activity which is not shown as separate output of the respective establishments in Table 2. The cost value of ancillary services is identified separately, however, in Table 3 as the total of intermediate consumption (17.9), consumption of fixed capital (4.8), compensation of employees (8.7) and net indirect taxes (0.3). These costs are balanced by a negative operating surplus (−31.7). It is not proposed to "externalize" the internal environmental protection activities within the SEEA in order to maintain close linkage with the SNA. For more comprehensive analyses of environmental expenditures and operations, ancillary activities could be externalized in supplementary tables.

The supply of products is disaggregated in Table 2 according to the three categories of natural growth products, external environmental protection services and other products only. A further breakdown of these categories needs to be developed, as far as possible in terms of the Central Product Classification (CPC) (United Nations, in prep.).

TABLE 1

SYSTEM FOR INTEGRATED ENVIRONMENTAL AND ECONOMIC ACCOUNTING (SEEA)

(Summary presentation)

					Tangible assets (Table 1.3)				
					Produced		Non-Produced		
					Except Natural	Natural (biota)	Natural		
Opening Stocks (Market Valuation)					991.3	83.1	1744.4		
					+(plus)				
					Capital Accumulation				
Use/Value added (Table 1.2)	Total	Domestic Production (Industries)	Final Consumption		Produced Assets		Non-Produced Natural Assets	Rest of the World	
			Households	Government	Except Natural	Natural (biota)		Exports/ Imports	Flow of Residuals
Use of products	591.9	224.0	175.0	42.5	68.0	1.4	7.3	73.7	
Gross Domestic Product (GDP)		293.4							
Consumption of fixed capital		26.3			-23.0	-3.3			
Net Domestic Product (NDP)		267.1			5.1	-0.9	-73.0		-4.7
Use of natural assets (ecological valuation)	-1.6	59.8	17.1	-5.0	5.1				
Environmental adjustment of final demand		22.2	-17.1						
Environmentally Adjusted Net Domestic Product (EDP)		185.1							
					+(plus)				
Supply/Origin (Table 1.1)									
Supply of products	591.9	517.4						74.5	
Origin of residuals	-1.6								-1.6
					+(plus)				
Adjustment of natural assets accumulation to market valuation						0.9	81.2		
Other volume changes (market valuation)					-25.3		22.8		
Revaluation due to market price changes					138.1	12.6	382.8		
					=(equals)				
Closing Stocks (Market Valuation)					1159.3	93.8	2165.5		

TABLE 2

SUPPLY/ORIGIN

		Domestic Production (industries)			Imports	
			Other industries			
	Total	External Environmental Protection Activities	Internal Environmental Protection Activities	Other Activities	Products	Residuals
(1) Supply of products (goods and services)	591.9	36.2		481.2	74.5	
(1.1) Natural growth products	40.7			38.2	2.5	
(1.2) External env. prot. Services	36.2	36.2				
(1.3) Other products 515.0	515			443	72	
(2) Origin of residuals -1.6	-1.6					-1.6
Σ Total supply [(1) + (2)]	590.3	36.2		481.2	74.5	-1.6

Natural growth products of agriculture, forestry and fishing (40.7) refer to those growth-based outputs that are controlled by human activities and can thus be considered as part of planned economic production. Natural growth in these products is treated as primary production which increases the stocks or fixed assets by the amount of growth taking place during the accounting period. On the other hand, those primary natural growth-based products that are largely harvested from the non-controlled natural environment (without human interference in the growth process, such as hunting, gathering of wild fruits, deep-sea fishing or the exploitation of tropical forests) are considered as either "free" inputs or, in case of "scarcities," as environmental depletion costs (see below, section 4) of the agriculture, forestry and fishing sectors. For example, in the case of free supply of fish to the fishing industry, the sector's output would not consist of live fish, but rather of fish landed and sold in the market-place.

4. Use and Value Added

The use/value added table (Table 3) shows the use of products and (man-made as well as natural) assets as inputs of the domestic production activities or as components of final demand (final consumption, capital accumulation, exports). These data are supplemented with information on the value added of the different production activities. The use table is an instrument to distribute the total supply of goods and services from the supply table to its various destinations. However, the supply of environmental assets is not displayed in the supply table, but is shown as negative entries in the natural non-produced assets column of capital accumulation. In comparison to the traditional framework of the SNA, the use of natural assets is shown in additional rows and the capital accumulation of non-produced natural assets in an additional column.

(a) *Use of Goods and Services*

The first block of rows in Table 3 presents the use of products (goods and services) by intermediate consumption of economic activities and final demand, as supplied from Table 2 (591.9). This corresponds to the traditional use table in the SNA. The sum of the gross value added (293.4), the conventional Gross Domestic Product (GDP), is shown explicitly in Table 3. Subtracting the consumption of fixed capital obtains the Net Domestic Product (NDP) (267.1).

As indicated in section 3, the supply of natural growth products (40.7) stems from "controlled" production processes of agriculture, forestry and fishing only. These products are used as inputs into different economic activities (23.0), exported (5.0), consumed by private households (11.3), or may increase fixed capital or stocks (1.4). Stock increase results from the growth in products which are not used in the same period. Stock decrease is shown where naturally grown products of a former period are used for intermediate or final purposes. The increase of fixed capital on the other hand represents a growth in the remaining biomass that is not intended to be used up in intermediate or final consumption,

TABLE 3
Use and Value Added

		Domestic Production (industries)				Final Demand							
								Net Capital Accumulation					
	Total	External Environmental Protection Activities	Other Industries		Subtotal Domestic Production	Final Consumption		Produced Assets		Non-Produced Assets	Exports		Sub-total Final Demand
			Internal Environmental Protection Activities	Other Activities		Households	Government	Except Natural	Natural (biota)		Products	Residuals	
(1) Use of products	591.9	15.9	17.9	190.2	224.0	175.0	42.5	68.0	1.4	7.3	73.7		367.9
(1.1) Natural growth products	40.7			23.0	23.0	11.3			1.4		5.0		17.7
(1.2) External envir. protection services	36.2			22.4	22.4	8.8	5.0						13.8
(1.3) Other products	515.0	15.9	17.9	144.8	178.6	154.9	37.5	68.0		7.3	68.7		336.4
Gross value added of industries [(9)-(1)]		20.3	-17.9	291.0	293.4								
													0.0
(2) Use of produced fixed assets (consumption of fixed capital)	0.0	1.3	4.8	20.2	26.3			-23.0	-3.3				-26.3
Net value added of industries [(9)-(1)-(2)]		19.0	-22.7	270.8	267.1								
(3) Use of natural assets (ecolog. valuation)	-1.6	6.3	4.6	48.9	59.8	17.1	-5.0	5.1	-0.9	-73.0		-4.7	-61.4
(3.1) Quantitative depletion	0.0	0.3	0.4	16.8	17.5	0.7			-0.9	-17.3			-17.5
(3.2) Degradation of land (except by residuals)	0.0	0.2		8.8	9.0	0.8				-9.8			-9.0
(3.3) Degradation by residuals	-1.6	5.8	4.2	23.3	33.3	15.6	-5.0	5.1		-45.9		-4.7	-34.9
Total use [(I) + (2) + (3)]	590.3	23.5	27.3	259.3	310.1	192.1	37.5	50.1	-2.8	-65.7	73.7	-4.7	280.2
(4) Environmental adjustment of final demand		1.8	2.1	18.3	22.2	-17.1		-5.1					-22.2
Env. adj. net value added (EDP) of industries		10.9	-29.4	203.6	185.1								
[(9)-(1) -(2)-(3)-(4)		10.9	-29.4	203.6	185.1								
or (5)+ (6)+(7)+ (8)]													
(5) Compensation of employees		13.0	8.7	72.0	93.7								0.0
(6) Indirect taxes minus subsidies		2.0	0.3	34.1	36.4								0.0
(7) Net operating surplus		4.0	-31.7	164.7	137.0								0.0
(8) Eco-margin [-(3) - (4)]		-8.1	-6.7	-67.2	-82.0								
(9) Total gross inputs/total total final demand [(1)+(2)+(3)+(4)+(5)+(6)+(7)+(8)]		36.2		481.2	517.4	175.0	37.5	45.0	-2.8	-65.7	73.7	-4.7	258.0

such as the trunks and branches of fruit trees or the breeding stock of livestock.[2]

External environmental protection services (36.2) are used for avoiding potential or restoring actual decreases in environmental quality. It is assumed in the numerical example that the environmental protection services of the government which are not sold on the market (government consumption: 5.0) are restoration activities whereas the other environmental protection activities (31.2) are avoidance activities and are bought by industries (22.4) and households (8.8). Environmental protection activities of the government for avoiding environmental degradation caused by its own production are assumed to be part of the internal environmental protection activities. Government environmental protection services services sold in the market are assumed to be intermediate consumption of industries or household consumption.

The other products (515.0) are used for intermediate consumption (178.6), final consumption (192.4), capital accumulation (75.3) and exports (68.7).

(b) *Use of Natural Assets*

Integrated environmental-economic accounting in the present framework focuses on the inclusion of costs, resulting from the quantitative depletion of natural resources and from the qualitative degradation of environmental quality by economic activities.

Depletion activities (at a total of 18.2) are shown in Table 3 to consist of depletion of natural assets by industries (17.5) and by households (0.7). As detailed in Table 5, they comprise the exploitation of natural resources such as sub-soil assets (mineral deposits) by mining and quarrying (−8.9), aquifers (−4.7) and biological assets (e.g. timber from tropical forests or fish stocks of inland and marine waters) by agriculture, forestry and fishing (−0.9, −3.7). The assumption is that scarcities in the availability of renewable (forest, fish, wildlife etc.) and cyclical (water) resources have been observed. Depletion costs are only estimated in these cases as far as the economic use of natural assets leads to imbalances in nature, i.e. if the depletion of biota exceeds the natural growth or the use of water exceeds replenishment of aquifers. The recording of corresponding negative amounts of tangible wealth reduction is discussed below in section 4d.

The other category of economic use of natural assets represents the environmental quality degradation of the environmental media of air, water and land by production and consumption activities. The degradation of land could be caused by improper agricultural practice (soil erosion, water logging, salinization), by excessive use for recreational purposes or by polluting the soil with wastes or waste-water. The main reason for degrading the quality of air and water is their use as a sink for residuals (wastes, pollutants) of economic activities. It has to be stressed that only the immediate influence on the environmental media is taken into account. The indirect effects by transboundary transport in the air or by transition from one environmental medium to another are not recorded in

[2]This treatment of natural growth processes in agriculture, forestry and fishing differs from the 1968 SNA recommendations, but may be adopted in the revised SNA.

the SEEA. These complex dynamics within the natural environment could be shown in supplementary data systems which should be linked with the SEEA. Furthermore, it should be noted that impacts of natural or man-made disasters are assumed not (or in some cases only indirectly) to be caused by economic use of environmental assets and are therefore excluded from the use/value added Table 3 but are included as a category (4) of asset volume changes in Table 5.

The net value of *degradation* is assumed to be equal to potential abatement (restoration) costs, required either to achieve the level of environmental quality at the beginning of the accounting period or at least a level specified by "official" environmental standards (Hueting, 1980). It is assumed that such standards reflect a technological solution to abating environmental quality degradation that can "reasonably" be expected to be applied by the different polluters. Obviously, such valuation does not measure actual environmental "damage" from pollution. A possible treatment of such welfare effects is discussed in section 5.

The environmental degradation is caused by production activities (9.0 plus 33.3), by consumption activities of households (0.8 plus 15.6), by man-made assets (5.1) and by imported residuals (1.6). Man-made assets have an effect on the natural environment by their residuals (e.g. scrapped machinery). A part of the environmental degradation is restored by government activities (−5.0). The remaining degrading impacts affect the domestic natural assets (−9.8, −45.9) and—as far as residuals are "exported"—the natural environment of the rest of the world (−4.7).

In Table 4, the value of the economic use (depletion as well as degradation) of domestic and non-domestic (foreign) natural assets and the corresponding impacts on the asset values are shown in a simplified balance sheet.

TABLE 4

ECONOMIC USE AND IMPACTS ON NATURAL ASSETS

Use of Natural Assets (environmental costs)		Impacts on Natural Assets (decrease of asset values)	
Domestic use		**Domestic environment**	
Depletion		Depletion	
industries	17.5	prod. natural assets	0.9
households	0.7	non-prod. natural assets	17.3
	18.2		18.2
Degradation		Degradation	
industries	42.3	non-prod. natural assets	55.7
households	16.4		
government	−5.0		
prod. assets	5.1		
Imports		**Environment of the rest of the world**	
Degradation	1.6	Degradation	4.7
	60.4		60.4
	78.6		78.6

(c) *Environmental Adjustments of the Value Added*

Deducting the imputed costs of natural asset use (environmental costs) from net value added leads to a new value-added concept, termed here "environmentally adjusted net value added." The environmental costs represent the hypothetical costs for maintaining the natural assets at the same level during the reporting period. This concept reflects a "strong" or "narrow" sustainability concept which implies that future generations should receive a natural environment with a quantitative and qualitative level being at least comparable with the present situation (Bartelmus, in preparation; Blades, 1989; Daly, in preparation; Pearce, Markandya, Barbier, 1989 and 1990; and Pezzey, 1989). The international discussion of the last years has proved that it is not sufficient to sustain a constant level of total (man-made as well as natural) capital, denying substitution possibilities between these capital categories ("broad" or "weak" sustainability concept). The uncertainty of long-term impacts of economic activities on the natural environment and the increasing knowledge about irreversible damages of natural balances (climate change, ozone layer depletion etc.) has led to a more cautious risk-conscious attitude towards overburdening the natural environment. From this point of view, it seems necessary to maintain the natural assets treating them as complementary to man-made capital. The strong sustainability concept thus applies not only to the case of environmental quality degradation, but also to the maintenance of "stocks" of natural resources. In the case of subsoil assets, this approach seem to be questionable because the strong sustainability concept would lead to non-use of the resources, possibly causing severe world-wide economic problems. Instead, the objective could be to maintain a long-term optimal depletion rate, considering that new finds could only retard the shrinkage of the stocks. It has been proposed that the sustainability concept should be weaker in this case, and it would be sufficient to balance a decrease of the subsoil assets with an increase of other types of assets (with preference for permanent or renewable natural assets) to sustain the same income level in the future (El Serafy, 1989; Daly, in preparation).

The maintenance cost approach used for valuing the economic use of natural assets corresponds to the methods of national accounting for estimating the use of man-made fixed assets. The user costs of these assets are estimated with the costs necessary to keep the man-made fixed capital intact, i.e. to maintain the level of the assets at the same level during the reporting period. These costs which are called "consumption of fixed capital" or "depreciation" are also used to compile the net capital formation of the man-made assets in the accounting period.

As far as the natural assets have the character of *fixed assets*, treating the maintenance costs of natural assets in the same way as the depreciation of man-made assets seems plausible. However, distinguishing between assets that bear characteristics of fixed assets and those that are more in the nature of an *inventory* or *stock* (in this case, decrease of assets in the national accounts is booked as intermediate consumption and not as depreciation) is problematic because natural assets may exhibit simultaneously economic and environmental functions (Hueting, 1980). For instance, a timber tract represents a stock resource, but has also an important role in cleaning the air and regulating water balances.

Furthermore, it serves as habitat of animals and as recreational area. From an ecological point of view, the environmental media, i.e. land, water and air as well as the ecosystems can be considered as fixed assets. The maintenance costs of these assets should therefore be treated as depreciation. Further discussion seems to be necessary in the case of subsoil assets. They mainly have the character of inventory stocks of nature. Their depletion could therefore be treated as intermediate consumption.[3] For sake of simplicity of the present framework, the value of the depletion of these assets is not shown separately from the other environmental costs, but is also treated as decrease of a fixed asset.

Whatever the treatment of the environmental costs, as depreciation of natural assets or as intermediate consumption, their deduction from gross output affects the calculation of net value added. The gross value added of the industries remains unchanged in the SEEA. The environmental adjustments of the net value added (−82.0) comprise the imputed environmental costs connected with domestic production (−59.8), household consumption activities (−17.1) and the use of man-made assets (−5.1). These adjustments are called eco-margin which is introduced explicitly in order to permit the identification of all components of value added (including operating surplus) according to the conventional SNA concepts.

Impacts of household activities and of man-made *assets* on environmental quality are taken into account for correcting the net value added despite the fact that the respective environmental costs are not directly associated with production activities. Regarding households, their polluting activities could be viewed as non-market production of goods and services which produces "jointly" residuals like wastes and pollutants. In this case, the net value added of the households' production would be diminished by the imputed environmental costs of the households. This is achieved by shifting these imputed values (17.1) from final consumption to the totals of domestic production. A similar correction is made with regard to the environmental impacts of man-made assets, comprising additional imputed costs of the asset owners (5.1). These costs refer e.g. to pollution caused by controlled landfill and to the residues of unrecycled man-made assets. It is theoretically possible to transfer these costs to the different industries. In this case, their net value added would directly be affected. This procedure has not been applied in the SEEA in order to show separately the environmental costs caused by current production and man-made capital use. The shift of the environmental costs of households and man-made assets to the columns of domestic production is shown in Table 3 in the row "environmental adjustments of final demand." Net value added is thus corrected only for the totals of environmental protection activities (1.8 and 2.1) and of other activities (18.3).

In Table 3 we also record the components of value added, consisting of the compensation of employees, indirect taxes net of subsidies, net operating surplus and environmental costs equal to item (8) "eco-margin." Use of SEEA thus permits the analysis of these components of value added for the environmental

[3]See El Serafy (1989). However, the total depreciation approach is advocated by Harrison (1989) and Repetto (1989).

protection activities of the different economic sectors. Indirect taxes and subsidies, charged or granted as part of environmental protection policies, will be identified separately in the SEEA, reflecting the application of polluter- and user-pays-principles at the micro-economic level. Macro-policy makers on the other hand, might be concerned with the assessment of employment devoted to "defensive" environmental protection activities (total "environmental" renumeration of employees of 21.7 as compared to a total of other wages and salaries of 72.0).

The net operating surplus of the different production activities has not been environmentally adjusted in Table 3. The additional environmental costs are balanced by introducing the eco-margin. The idea is to facilitate the unequivocal linkage of the production accounts of the SEEA with the income accounts of the conventional SNA. Another possible presentation of the operating surplus could show an environmentally adjusted net operating surplus, but extend the table at the same time for identifying explicitly the non-adjusted gross and net operating surplus:

Environmentally adjusted net operating surplus	55.0
= Gross operating surplus	163.3
− Consumption of fixed capital	26.3
− Eco-margin	82.0

The total of the environmentally adjusted net value added is called Environmentally Adjusted Net Domestic Product or, short, Eco Domestic Product (EDP). EDP could be derived from GDP as follows:

Gross Domestic product (GDP)	293.4
− Consumption of fixed capital	26.3
= Net Domestic Product (NDP)	267.1
− (Imputed) environmental costs	82.0
= Eco Domestic Product (EDP)	185.1

(d) *Final Demand*

Final demand consists of final consumption, net capital accumulation and exports. Import and export flows are only slightly modified for environmental accounting. However, significant alterations are proposed for both final consumption and net capital accumulation to allow corrections of net value added while heeding the principle of accounting identities (see below, section 4e).

Imports and exports include flows of wastes which are not marketed, but transported to/from a foreign country or to the open sea. They represent either a degradation of the foreign natural media by exporting domestic residuals or of domestic media by importing residuals. They are estimated as negative values—costs for avoiding or restoring environmental quality degradation (exports: −4.7, imports: −1.6). The imports of residuals reduce the total value of imports and of the domestic natural assets (negative item in the column of net accumulation of natural assets). The exports of residuals lead to an increase of the imputed environmental costs of the exporting industry which implies a reduced environmentally adjusted net value added of the exporting economic units and a reduction of the total value of the exports. Transboundary flows of residuals of the economic activities which are not transported by man but by environmental media (e.g. water, air), are recorded as degradation of the environmental media which directly receive the residuals. Their final destination is not taken into account.

The conventional final consumption of households (175.0) remains unchanged in the SEEA. The additional (imputed) environmental costs (17.1) which would have been necessary to avoid or which need to be incurred in order to restore a degradation of environmental quality by household activities (recreational use of land 0.8, pollution 15.6), or which represent the costs of depleting natural resources (firewood consumption 0.7) are shifted to domestic production.

The fully elaborated SEEA will comprise a further breakdown of the final consumption of households by environmentally oriented functions for identifying e.g. the environmental protection expenditures of the households and the expenditures required to compensate for the damages caused by environmental deterioration (health expenditures etc.).

Final *consumption of government* (42.5) is corrected by the environmental protection expenditures (−5.0) which are non-marketed and which are undertaken to avoid or restore a decrease of environmental quality caused by other economic units. These expenditures have the characteristics of an investment in environmental quality. Its value is shifted from the government final consumption to the capital accumulation of natural assets and diminishes the degradation of the natural assets which would have occurred if no restoration activities had taken place. The environmental protection activities of the government for own purposes (internal activities) and the additional (imputed) environmental costs of the government production activities are already recorded in the columns of domestic production. It is therefore not necessary to extend the concept of final consumption of government in the SEEA by taking into account imputed environmental costs.

The section on the *net accumulation of tangible wealth* in Table 3 differs considerably from the traditional incorporation of capital formation in a use table. The presentation of this part of final demand in Table 3 is limited to an asset classification by only three types of asset: produced biological (natural) assets, other produced assets, non-produced natural assets. A further breakdown of the capital accumulation is given in Table 5 which shows complete asset balances for the different types of assets. The following comments refer especially to the disaggregated version in Table 5. The capital accumulation concept of Table 3 corresponds to the (traditional) capital formation (item 2 of Table 5) and to the ecological valuation of volume changes of natural assets due to economic use (item 3.1 of Table 5). The valuation problem (market versus ecological valuation) and the other items of Table 5 are discussed in section 5.

In the SEEA, the asset boundary has been extended for including all natural assets which are actually or potentially used by economic activities or which could be affected by the residuals of economic production and consumption

TABLE 5
Asset Balances of Net Tangible Assets
Monetary units

		Total	Produced Assets		Non-produced Natural Assets					
			Produced Assets (except biological)	Produced Biological Assets	Non-produced Biological Assets	Land (landscape, ecosystems)		Sub-soil Assets	Water	Air
						Cultivated, etc.	Uncultivated			
(1)	Opening stocks (market values)	2818.8	991.3	83.1	65.4	1366.7	50.4	261.9		
(2)	Net capital formation (use of products, market values)	50.4	45.0	-1.9		4.6		2.7		
	(2.1 Gross capital formation	76.7	68.0	1.4		4.6		2.7		
	(2.2 Consumption of fixed capital	-26.3	-23.0	-3.3						
(3)	Volume change of natural assets due to economic use (market values)	36.0			-2.1	23.3	-5.0	19.8	0.0	0.0
	(3.1 Ecological valuation									
	(3.1.1) Quantitative depletion	-18.2		-0.9	-3.7			-8.9	-4.7	
	(3.1.2) Degradation of land (except by residuals)	-9.8			-7.7	-2.1				
	(3.1.3) Degradation by residuals	-45.9				-9.5	-3.1		-12.9	-20.4
	(3.2 Adjustment due to market valuation									
	(3.2.1) Quantitative depletion	8.1		0.9	1.6			0.9	4.7	
	(3.2.2) Land use (except by residuals)	35.2				33.1	2.1			
	(3.2.3) Degradation by residuals	38.8				4.0	1.5		12.9	20.4
	(3.3 Other volume changes (change of land use, new finds, new estimates etc.)	27.8				3.4	-3.4	27.8		
(4)	Volume change by natural or multiple causes (market values)	-30.3	-25.3		1.3	-4.3	-2.0			
(5)	Revaluation due to market price changes	533.5	138.1	12.6	11.1	331.0	11.8	28.9		
(6)	Closing stocks (market values) [(1)+(2)+(3)+(4)+(5)]	3408.4	1149.1	93.8	75.7	1721.3	55.2	313.3		

processes. The extended asset concept comprises the following types of assets:

Produced assets
- Man-made assets (non-biological such as machinery and equipment, stocks of non-biological products)
- Natural assets produced by agriculture, forestry and fishing (fixed assets and inventory stocks)

Non-produced natural wild assets
- Wild biological assets
- Land (cultivated and uncultivated)
- Subsoil assets (developed and undeveloped proven reserves)
- Water (stored and unstored)
- Air

This classification distinguishes in particular between assets which are (economically) produced or not and which are man-made or natural. These two criteria are not identical because the (economically) produced biota are both produced and natural. In this case, the produced biota should only be subsumed under natural assets as far as they are living. A further breakdown is possible according to the degree of human influence on the natural environment (e.g. cultivated-uncultivated land, developed-undeveloped subsoil assets).

In Table 3, the net capital accumulation of *produced assets* is recorded mainly according to the conventional concepts of the SNA (gross capital formation: 68.0 and 1.4, consumption of fixed capital: −23.0 and −3.3). Only two minor deviations should be mentioned: The residuals of the produced assets which are loaded on the natural environment (e.g. scraps, pollution of controlled landfill) are valued with their avoidance costs (5.1) and shown in addition to the net capital formation. In a second step, these imputed costs are shifted to the industries of the responsible activities or, alternatively, to the industries as a whole [via the environmental adjustment row (4)]. In the case of produced biological assets, it might be necessary to estimate additional depletion costs (−0.9) if the economic activities of agriculture, forestry and fishing disturb the natural balances, e.g. if the amount of cut wood exceeds the natural growth and destroys the ecosystems of cultivated forests. In this case, the sustainability principle should be applied, and avoidance or restoration costs could be calculated.

The imputed depletion costs of *wild biota* (in Table 5: −3.7) are estimated only if depletion by economic activities (e.g. hunting, ocean fishing) and natural growth are not balanced. Depletion costs are thus estimated if depletion exceeds natural growth. The discussion of valuation of net depletion in this case has not been conclusive. One possible approach could be to value net depletion as the gross value added generated by the depleting activity. This would show value added foregone if the net depletion had been avoided. Another approach could be to assess the costs for compensating projects to restore the natural balances.

Net capital accumulation of *land* refers to the impacts of economic land use. The costs of developing land are treated in the conventional SNA as capital formation which normally leads—from an economic point of view—to an improvement of land quality and to increasing market values (in Table 5: 4.6).

From an ecological point of view, increasing economic use of land could cause a qualitative degradation of the land and the terrestrial ecosystems. The main reasons are restructuring (further economic development of cultivated land, cultivation of uncultivated land), intensive agricultural use (soil erosion etc.), recreational use (disturbing ecosystems) and use as a sink for residuals (such as pesticides and the pollution of controlled and uncontrolled landfill). In Table 5 degradation by residuals (−9.5 and −3.1) and by other economic activities (−7.7 and −2.1) are distinguished.

The degradation of land is valued as the cost to avoid (or at least mitigate) the negative impacts of economic activities or to restore the degraded areas, with a view to maintaining the terrestrial ecosystems in their present state. This valuation concept might differ widely from the market valuation. Changes in economic land use will often increase the market value of land, but at the same time could imply a decrease of the ecological quality of land.

Subsoil assets comprise the proven reserves of fossil and mineral assets. Proven reserves normally have to meet three criteria: high probability of existence (95 percent), exploitability with existing techniques and positive net return, i.e. the market price exceeds exploitation costs (Martinez *et al.*, 1987). Subsoil assets can be undeveloped or developed (established mines and other exploration facilities). The costs for developing subsoil assets (e.g. by exploration activities) have to be treated according to the conventional SNA as capital formation (Table 5: 2.7). The valuation of the depletion of subsoil assets has to reflect the future scarcity of the assets. Exploitation is mainly an economic and not an ecological problem because the immediate impacts on natural balances are usually low, with the notable exception of surface mining. The indirect impacts of subsoil depletion (e.g. losses of crude oil during transportation, pollution connected with energy consumption) are registered independently from the valuation of assets as environmental degradation from polluting economic activities.

Various methods have been proposed to value subsoil asset depletion (see Ward, 1982). Several authors suggest the use of the net operating surplus of the exploiting industry or a part of it. The proposal of El Serafy (1989) seems to be an approach tailored to the concept of sustainability. The idea is to estimate the depletion costs as the amount of money which should be invested to achieve a long-term constant flow of income, even after complete exploitation of the resources. This rule implies a substitution of the use of subsoil assets by other types of income generating activities and corresponds to a broad sustainability concept. The decrease of subsoil assets could be balanced, e.g. by increasing renewable (biological) assets or by the development of solar and wind energy sources instead of coal or crude oil. The value of subsoil depletion amounts to 8.9, in Table 5.

The economic use of *water* could lead to increasing scarcity (depletion) or to decreasing quality (degradation by residuals). Increasing scarcity of water will be observed if the economic abstraction exceeds the average natural inflow of water during the accounting period. In this case, net depletion could be valued as the value added, or part of it, generated by additional water use of the abstracting industries (−4.7). This value could represent the avoidance costs as in the case of wild biota. Further discussion is necessary to develop a generally accepted valuation method for water depletion. The degradation of water by residuals is valued as its avoidance (or restoration) costs (−12.9).

As described above (section 4b), the value of the degradation of *air* is its avoidance costs (−20.4).

In section 5 below, we describe a comprehensive system of balance sheets of tangible assets, including changes in volume not accounted for by capital accumulation in the use/value added table (e.g. effects of natural and other disasters).

(e) *Accounting Identities*

The national accounting identities between the totals of environmentally adjusted value added (plus imports) and final demand are maintained in the use/value added Table 3 by treating capital accumulation of natural assets as part of final demand. In Table 6 we show the transition from the conventional aggregates, according to the SNA, to the environmentally adjusted aggregates of the SEEA by using the numerical example (cf. Tables 2 and 3).

TABLE 6

ACCOUNTING IDENTITIES

Primary Inputs (value added, imports)		Final Demand (domestic, exports)	
Gross value added (Gross domestic product)	293.4	Domestic final demand (SNA concept)	294.2
− Use of produced assets (consumption of fixed capital)	26.3	− Consumption of fixed capital	26.3
− Use of natural assets for current production	59.8	− Government restoration costs	5.0
Environmental adjustment of final demand	22.2	+ Net capital accumulation of natural assets (−78.9 + 5.0)	−73.9
Environmentally adjusted net value added (Environmentally adjusted Net Domestic Product)	185.1	Environmentally adjusted domestic final demand	189.0
+ Import of products	74.5	+ Export of products	73.7
+ Import of residuals	−1.6	+ Export of residuals	−4.7
Environmentally adjusted primary inputs	258.0	Environmentally adjusted final demand	258.0

As already explained, the gross value added (293.4) is corrected by the consumption of fixed produced capital (26.3) and by the user costs of natural assets (current production 59.8, households 17.1, residuals of man-made assets 5.1). This obtains the environmentally adjusted net value added (185.1). The concept of imports is extended to additionally include imports of residuals (74.5 and −1.6). The total of the environmentally adjusted primary inputs (value added plus imports) is 258.0.

Domestic final demand (294.2) is corrected by the consumption of fixed produced assets (26.3) to achieve a net concept. The environmental restoration costs of the government (5.0) are treated as an increase of the value of natural assets and therefore reflected in the capital accumulation of natural assets. The value of depletion and degradation of natural assets by economic activities would have been −78.9 without government restoration activities. Taking these activities into account, the net capital accumulation amounts to −73.9 (−73.0 plus −0.9). The environmentally adjusted figures of total final demand (258.0) comprise the adjusted domestic final demand (189.0) and the export of products (73.7) and residuals (−4.7). This total is equal to the total of primary inputs.

5. Asset Balances of Tangible Wealth

As illustrated in Table 1, the section of the use/value added table on tangible wealth accumulation can also be viewed as an integral part of the asset balances. This is indicated in Table 1 by plus (+) and equal (=) signs, inserted between the four elements of the asset balances. This illustration shows an accounting identity between the closing stocks and the sum of opening stocks, net capital accumulation, adjustment of natural assets accumulation to market valuation, other volume changes and revaluation due to market price changes. This identity holds for man-made assets (produced, not natural):

$$1{,}149.1 = 991.3 + (68.0 - 23.0 + 5.1 - 5.1) + (-25.3 + 138.1),$$

for the (economically) produced natural assets:

$$93.8 = 83.1 + (1.4 - 3.3 - 0.9) + (0.9 + 12.6),$$

and for those non-produced natural assets which are used or affected by economic activities:

$$2{,}165.5 = 1{,}744.4 + (7.3 - 73.0) + (81.2 + 22.8 + 382.8).$$

The asset balance sheets are further elaborated on in Table 5. The asset classification has already been described in Section 4d. The volume and price changes of the assets during the reporting period are further disaggregated in Table 5, consisting of:

(2) Net capital formation (use of products);
(3) Volume change of natural assets due to economic use;
  (3.1) Ecological valuation,
  (3.2) Adjustment due to market valuation,
  (3.3) Other volume changes (market valuation),
(4) Volume change by natural or multiple causes;
(5) Revaluation due to market price changes.

The volume changes (2) and (3.1) reflect the net capital accumulation described in the use/value added table (Table 3 and Table 1.2), the adjustment due to market valuation (3.2) corresponds to the "adjustment of natural assets accumulation to market valuation" in Table 1.3. The other volume changes due to economic use (3.3) and the volume change by natural or multiple causes (4) are summarized under "other volume changes" in Table 1.3. The revaluation due to market price changes is presented in both tables under the same name.

The design of the asset balance sheets aims at introducing environmental aspects in the national stock accounts without disrupting the concepts of the conventional SNA balance sheets. As recommended in the International Guidelines on Balance-Sheet and Reconciliation Accounts (United Nations, 1977) and in chapter XI of the preliminary draft of the revised SNA (United Nations, 1990), the *opening* and *closing stocks* are valued at market prices or have values derived from market prices. Direct market valuation could be applied if the assets are marketed (some produced fixed assets like cars, inventory stocks of products, land). Indirect market valuation uses the net value concept (replacement costs minus cumulated depreciation) or tries—in the case of depletable natural assets like wild biota, subsoil assets or water—to estimate the assets by the discounted value of future net returns (future market prices minus all exploitation costs including a normal rent of capital).[4] It should be stressed that the SEEA does not aim at a complete market valuation of the non-produced natural assets. The market valuation should be limited to natural assets which are regularly depleted for market purposes (e.g. ocean fish, tropical wood and subsoil assets) or to assets which are directly marketed (uncultivated land in exceptional cases). The opening and closing stocks of the other non-produced natural assets have a market value of zero. In these cases, their volume changes are valued only if they are affected by economic activities.

Market valuation has also been applied in general for the *volume* and *price changes* during the accounting period. Net capital formation of produced assets [item (2) in Table 5] reflects the volume changes described in the conventional SNA framework. Some of the other volume changes of assets caused by economic, natural, non-economic and multiple (combination of these causes) activities and events [items (3) and (4) in Table 5] which had been part of the reconciliation accounts in the 1968 version of the SNA, will be integrated into the accumulation accounts which explain the changes in the balance sheets of the revised SNA at market values. Opening and closing stocks of these assets are also measured at market values. The transition to the level of the market values at the end of the accounting period (closing assets) is shown as item (5) in Table 5 (revaluation due to market price changes).

The connection between the SEEA and the conventional assets balance sheets is introduced by a breakdown of the volume change of natural assets due to economic use (item (3) of Table 5). As far as the economic use affects the natural balances and leads to a decrease of the value of the natural assets from an ecological point of view, the avoidance or restoration costs are estimated for maintaining the same qualitative and quantitative level of natural capital during the accounting period. These values are introduced in the extended use/value added table of the SEEA (see Table 3). This ecologically oriented valuation does not necessarily correspond to the market values due to the respective economic use. Therefore, an adjustment item is introduced which allows the transition to the market valuation of the asset balance sheets [item (3.2) of Table 5]. Volume

[4]The normal rent of capital refers to the produced assets which have been used for the exploitation of natural assets (e.g. trawlers for fishing and drilling instruments).

changes due to economic activities which do not directly deplete or degrade the natural assets (changes in land use, discoveries, etc.) are separately recorded and have market values [item (3.3) of Table 5]. In the SEEA, the analysis of the volume changes of the natural assets is focused on the economic uses. Therefore, no ecologically orientated valuation of the volume changes of natural assets due to natural or multiple causes (like wars and disasters) is applied, but market values of the volume changes are given.

It is not possible to describe the volume changes of the different types of assets in this overview article in detail. More detail on the extended asset balance sheets will be given in the SNA Handbook on Environmental Accounting. The following limited observations are thus only to facilitate a better understanding of the general scope and coverage of environmental assets in the SEEA.

The consumption of fixed capital (fixed produced assets) comprises only insurable risks of premature losses. Further losses by war or natural disasters are recorded under item (4) (−25.3).

The asset balances of the *biological assets* are relatively complicated because of the different concepts for describing the volume changes of produced and non-produced biota at market values. The natural growth of produced biota is treated as economic production (and gross capital formation) whereas the natural growth of non-produced biota is, as far as market values are associated to them, part of "(4) Volume change by natural or multiple causes" of Table 5. The depletion (due to economic use) of the produced biota is shown as decrease of stocks or consumption of fixed capital [item (2) of Table 5], whereas the depletion of non-produced biota is indicated under item (3). This different treatment implies that the net growth (natural growth minus depletion) of produced biota at market values is shown as net capital formation under item (2) ($-1.9 = 1.4 - 3.3$), while the net growth of non-produced biota equals the difference of the values under item (4) and item (3) ($-0.8 = 1.3 - 2.1$).

The ecological value of depleting produced biota (−0.9) reflects the ecological consequences of depleting cultivated biota beyond the economic use of these assets. If, for example, the wood of timber tracts is cut, the natural balance of forests could be disturbed as far as depletion exceeds natural growth. The necessary "ecological" costs could be estimated by the costs of compensating projects or by the additional value added generated by the net depletion (gross depletion minus natural growth). Further considerations are necessary to avoid double-counting if the depletion of produced biota exceeds their natural growth. In this case, the ecological valuation has to take into account that the (negative) net growth has already been valued at market values as (negative) capital formation.

The depletion of wild biota (−3.7) could be valued in ecological terms in a similar way. The natural balances could only be maintained if depletion and natural growth are balanced. That is, if the net depletion is positive avoidance or restoration costing refers to a reduction of the production (hunting, harvesting etc.) and a decrease of the corresponding value added, and the loss of ecological functions of the resource.

The valuation of the quality changes of *land* caused by economic activities might produce completely opposite results, depending on the economic or environmental point of view. Restructuring and development of land are normally connected with increasing market values, whereas their ecological effects could decrease the land values under environmental aspects. The development costs (4.6) are shown as capital formation. They reflect, together with the market value of the volume changes due to economic use (23.3, −5.0), the market value of all quantitative and qualitative volume changes of land caused by the different economic activities. The quantitative aspects of land use (changes in land use) are described under item (3.3) of Table 5 (3.5, −3.4). The qualitative component is shown under ecological aspects first (−7.7, −9.5, −2.1, −3.1) and adjusted to market valuation in a second step (33.1, 4.0, 2.1, 1.5). The qualitative changes do not only comprise the results of restructuring and development, but also excessive economic use, e.g. for agricultural purposes (often connected with soil erosion) and for recreation. Furthermore, the degradation by residuals is taken into account.

The ecological valuation of land degradation raises difficult estimation problems. In principle, the adequate avoidance or restoration costs to maintain the same level of land quality has to be estimated. Avoidance costs could comprise the decrease of value added in case of reducing excessive land use. Restoration costs could be the costs of compensation projects.

The opening and closing stocks of *subsoil assets* (proven reserves: developed and undeveloped) are valued with the discounted value of future net returns (i.e. revenues minus exploitation costs: 261.9, 313.3). New discoveries and changes in the economic conditions of exploitation which lead to new estimates of the proven reserves, are shown under item (3.3) of Table 5 (27.8). The exploitation costs do not contain the exploration costs (2.7) because they have already been included under item (2) of Table 5. The extraction (depletion) of these assets is estimated at "ecological" values (−8.9) and in a second step adjusted to market values (−8.9+0.9 = −8.0). These market values reflect the net prices of the depleted assets (current market price minus exploitation costs). The ecological valuation could comprise the costs for maintaining the level of natural capital (compensating projects to develop renewable or permanent assets) or of the total capital (man-made and natural).

The stock of *water* has normally no market values. Exceptions are stored water for drinking or irrigation purposes. The depletion of water is valued from an ecological point of view only if the average water stock is affected. This net depletion is valued with its avoidance costs [costs of reducing water use, e.g. by reducing agricultural production (−4.7)]. The avoidance cost approach can also be applied for valuing water degradation by residuals (−12.9).

The *air*, as a natural asset, has no market value. Therefore, the value of opening and closing stock is zero. For balancing the value of degradation by residuals (−20.4), a corresponding positive item has been introduced as an adjustment to market valuation.

6. Welfare-Oriented Measures of the Economic Use of the Environment

The concept of sustainability used in this paper is cost-oriented rather than welfare-oriented.[5] It reflects cost estimates which would be necessary to avoid, restore or replace decreases of environmental quantities and qualities during the reference period. Such an approach would normally suggest a greater effort at protecting the environment, as compared to estimating an economically optimal level of pollution. Optimality would require a balance of marginal costs of protection activities and of the (discounted) flows of marginal future environmental damages avoided. Because of underestimation, uncertainty and undervaluation (high discounting) of future damage, the optimality criterion will almost certainly present an amount of environmental deterioration which might be optimal from a micro-economic point of view, but not from a social point of view. In view of the uncertainties related to individual (marginal) evaluation and of prevailing societal and international concerns over long-term threats to critical life-support systems, the cautious concept of sustainability implicit in the cost values of the present framework, i.e. the maintenance (non-decrease) of environmental quality, appears to be a realistic approach. Under this aspect, the present cost approach also reflects (social) welfare aspects in its valuation of environmental degradation. Theoretical considerations, recently presented by Pearce, Markandya, Barbier (1990, especially p. 9), seem to support this approach.

Measurement and valuation problems of estimating the consequential damage (welfare losses) caused by environmental degradation are formidable. It is also difficult to associate unequivocally particular pollutants with health and welfare effects (for example, health damage caused by air pollution). One approach proposed to assess damage costs is to measure actual expenditures required for the elimination of the damage (Uno, 1989). Such expenditures could be shown separately in the SEEA as possible deductions under welfare aspects (Leipert, 1989). Another approach is to directly estimate health and welfare losses, including the impairment of recreational functions or aesthetic and ethical aspects of the environment. Some of these losses have been estimated by using the willingness-to-pay approach as an approximation of individual ("revealed") preferences or by other methods of contingent valuation (see OECD, 1989). Once comprehensive estimates of the value of damages become available, research projects could be undertaken to associate them with the polluting sectors. In this case, separate accounts should be established, which would allow a comparison of the actual and hypothetical avoidance cost on one hand and of the actual and imputed damage costs on the other. These comparisons would facilitate macroeconomic cost-benefit analyses, as proposed for example by Peskin (1989). Such additional accounts would permit further modifications of the components of final demand for the derivation of welfare-oriented measures (Bartelmus, 1987). In the SNA Handbook on Integrated Environmental and Economic Accounting, an approach will be discussed which could be derived from the cost-oriented measures of the SEEA by extending not only the asset boundary, but also the production boundary. This implies the introduction of the concept of environmental services "produced" by nature (see Peskin, 1989; Schäfer and Stahmer, 1989; and Stahmer, 1990).

---

[5]The question of welfare-oriented measures in national accounts is discussed in Drechsler (1976) and United Nations (1977a). The limits of accounting approaches to assets, the sustainability of economic growth, and possibilities of modeling the "feasability" of development programs are discussed in Bertelmus (in preparation).

7. Application of the Framework: A Desk Study of Country X

The environmentally modified concepts developed in the framework should stimulate alternative economic analyses and policies, based on an integrated assessment of environment-economy relationships. One aspect of such analysis is to focus on income available for spending on final consumption and new investments. Due to the consideration of environmental costs of production, environmentally adjusted income would generally be lower than income derived in traditional accounting. This *welfare* aspect of environmental accounting has received most of the attention in environmental studies of income and expenditure.

On the other hand, production cost and tangible asset and resource requirements of production reflect a *productivity* aspect of economic performance and environmental-economic analysis. Environmental accounting may result in values of value added generated and tangible assets used in each sector, that are different from the values of income and capital in traditional accounting. The reasons are the inclusion of cost due to environmental uses and of non-produced natural assets in broader concepts of cost and capital, respectively. Changed relations between value added and economic assets used in production might well lead to considerable re-assessment of the rentability and productivity of economic sectors from an environmental (accounting) point of view.

(a) *Economic and Environmental Features of Country X*

This analysis is done on the basis of an illustrative database, developed for the clarification of the above-described environmental accounting concepts and procedures. Only part of this database is reflected in Tables 1 to 5 above. The data describe the economic and environmental features of a realistic, but fictitious country "X" and are thus to a large extent fictive. There is however a basic core of data which was taken from the national accounts of an existing country. These core data include GDP by activity and expenditure categories, compensation of employees, indirect taxes (net of subsidies), operating surplus, output, intermediate consumption, capital formation and the consumption of fixed capital.

All other national accounts data, included in the framework, are elaborated on the basis of assumptions about the type of country, the circumstances under which traditional GDP is being generated in production, the effects of production on the environment, and environmental protection and response carried out by government, enterprises and individuals. These assumptions permit the breakdown of aggregate economic data which are part of conventional national accounts, but which could not be compiled from original data sets. Further assumptions about the environmental conditions of the country were made to obtain environmental data for calculating the environmentally adjusted concepts of income and expenditure.

The economic and environmental features of the fictive country X, as reflected in the data and assumptions, are described in the following. The country is a developing country with oil resources, agricultural production, exploitation of timber resources, and fishing activities in rivers, lakes and ocean.

(i) Tangible Wealth

Fixed assets consist of buildings, machinery and equipment, roads and other public structures, and also of livestock for breeding, draught and dairy, trees in orchards and grapevines in vineyards. As regards land, the assumption is that it is used mainly in agriculture for the cultivation of crops and rearing of livestock, in other services for dwellings and office buildings, and that it is owned as infrastructure by government (roads, dams and other structures).

(ii) Environmental Protection

Environmental protection activities are carried out in the country by all sectors. They are concentrated, however, in three sectors which are selling protection services: (a) "other services" which provide private waste disposal services, environmental consulting and recycling; (b) government, which provides sanitation services, and (c) trade and transport, which transports wastes to dumping areas and treatment and recycling plants. The environmental protection (sanitation) services offered by the government are sold to a very limited extent, the rest is assumed to be used by the government itself. The value of the cleaning activity is assumed to be equal to the cost (5.0). Households also purchase environmental protection services. These purchases (8.8) are presented as expenditure in the column of "household final consumption" in Table 3.

(iii) Mining Exploitation

The value of the mineral deposits of the country consists in particular of the value of oil reserves (opening stock in Table 5: 261.9). New deposits of oil were found as a result of exploration activities. This is presented in Table 5 as "Other volume change" (27.8), exceeding the amount of depletion ($-8.9+0.9$).

(iv) Natural Growth

The country has an important agricultural sector, a fishing industry which operates in rivers, lakes and the ocean, and timber tracts where wood is cut and replanted in a controlled exploitation activity. There are also minor wood collection activities in rural areas which are not controlled by any permits.

(v) Natural Disasters

During the period of accounting, the country suffered from a major earthquake which destroyed some of its infrastructure, particularly affecting roads owned by the government, machinery and equipment in the manufacturing sector, and dwellings and other buildings that are recorded as capital of the other service sector. The total value of destruction is included in Table 5 ("Volume change by natural or multiple causes:" −25.3).

(vi) Pollution

Pollution effects as a result of economic activities in the country are recorded in Table 3 in the row corresponding to qualitative degradation of land, water and air by residuals. In the case of air and water pollution, it is assumed that not only the domestic air and water were affected, but also those of neighboring countries (−4.7). Private households also cause pollution, which is assumed to consist of the effects of accumulated and illegally discharged wastes. The cost of this pollution (15.6) is reflected in the intersection of qualitative degradation by residuals and the household consumption.

(vii) Conversion of Tropical Forest to Commercial Use

Tropical rain forests are being converted to a limited extent to land for agriculture, urbanization and industrial development. This is recorded as change of land use in Table 5 (±3.4).

(b) *Comparative Analysis of the Economic Conditions of Country X, based on National and Environmental Accounts*

In Tables 7 and 8 we compare aggregates and indicators from traditional accounting with corresponding ones in environmental accounting. NDP is the main concept of national accounting and EDP is used as the environmental accounting alternative.

TABLE 7

ANALYTICAL MEASURES IN TRADITIONAL AND ENVIRONMENTAL ACCOUNTING: RESOURCES AND USES

Macro-aggregates	Based on: NDP (1)	EDP (2)	Percentage Difference (2)−(1)/NDP
Income/expenditure	267.1	185.1	−31
Final consumption	217.5	212.5	−2
% of final domestic uses	81	112	
Capital formation (accumulation) net	50.4	−23.5	−28
% of final domestic uses	19	−12	
Exports	73.7	69.0	−2
Minus: imports	74.5	72.9	−1

(i) Income and Expenditure

The income and expenditure analysis is done on the basis of the national accounts identity between income on one hand and domestic expenditure (final consumption, investment) plus exports minus imports on the other. The income and expenditure aggregates, based on Tables 2 and 3, are presented in Table 7.

In the case of traditional national accounting (column 1 of Table 7) the income concept is NDP, and the expenditure concepts are final consumption, net capital formation and exports minus imports. In the case of environmental acccounting (column 2), NDP is replaced by EDP, consumption by environmentally adjusted consumption and net capital formation by net capital accumulation. Environmentally adjusted consumption is derived from final consumption by deducting the improvement in the environment which is assumed to be equal to the government's net (accounting for clean-up of its own pollution) expenditure for environmental protection (5.0). Environmentally adjusted capital accumulation (−23.5) is arrived at by deducting from net capital formation (50.4) total environmental uses for all (produced and non-produced) asset categories (73.9).

The two sets of aggregate data present a very different picture of the economic situation of the country. Income is reduced drastically between NDP and EDP from 267.1 to 185.1, which represents a 31 percent reduction. Most of this reduction is caused by the modifications of the concept of capital formation to obtain a new concept of capital accumulation. Net capital formation changes from being positive (50.4) to a negative net capital accumulation of −23.5, which constitutes a reduction of income of 28 percent. The remaining 3 percent of the total reduction of GDP are explained by the difference between final consumption and environmentally adjusted consumption (−2 percent) and the decrease of exports minus imports (−1 percent) (see column 3, Table 7).

According to conventional national accounting, the country's domestic expenditure presents a healthy picture of capital formation of 19 percent of total expenditure. Environmental accounting indicates, however, that capital accumulation has been negative. The main factor explaining this result is the inclusion of non-produced assets into the asset boundary: the depletion of natural assets reduces the capital formation by a value of 18.2; a further reduction (−55.7, see Table 3) results from the degradation of land, water and air.

(ii) Income, Output and Capital

Production-related changes in environmental accounting, as compared to traditional national accounting, are elaborated on in the three sections of Table 8. In Table 8 we use figures for specific industries which are not shown in Tables 2 and 3 above, but which represent disaggregated (by economic sectors) figures of these tables. Section (i) of the table shows that there is a reduction in value added for the economy as a whole with EDP amounting to 69 percent of NDP. The impact differs from sector to sector, however. The largest reductions are in mining (52 percent) and agriculture (49 percent). In manufacturing, the reduction is 22 percent. Trade and transport also show a reduction of 14 percent due to the environmental cost of traffic pollution. All other sectors have lower reductions.

These differences are also reflected in the ratios of value added over output under NDP and EDP calculations in section (i) of Table 8. For the economy as a whole, the ratio falls from 52 percent to 36 percent. The largest drop is in agriculture from 77 percent to 39 percent, followed by mining (44 percent to 21 percent), trade and transport (61 percent to 52 percent), and manufacturing (34 percent to 27 percent). Other sectors show much lower reductions. Consequently, there are changes in the order of sector contributions to EDP as compared to NDP. Trade and transport is the largest contributor to both EDP and NDP. Manufacturing is the second largest and other services the third largest contributor to NDP. For EDP, however, this order is inverted. The weight of agriculture and mining in the economy is decreased whereas construction and government services increase in importance.

The other element of production cost, the use of economic wealth, is also affected by differences in coverage between traditional national accounting and environmental accounting. As shown in section (ii) of Table 8, produced assets, which are the capital element in national accounting, for the economy as a whole amount to only 38 percent of the total value of capital used, if non-produced assets are taken into account. For individual sectors the differences in coverage

TABLE 8

ANALYTICAL MEASURES IN TRADITIONAL AND ENVIRONMENTAL ACCOUNTING: INCOME, OUTPUT AND CAPITAL

(in percents)

Analytical Measures	Total	Industries: Agriculture	Mining	Manufacturing	Electricity, Gas, Water	Construction	Trade and Transport	Other Services	Government Services	Environmental Adjustments
(i)										
EDP as percentage of NDP	69	51	48	78	97	90	86	91	96	
Value added as percentage of output, based on:										
NDP	52	77	44	34	36	45	61	62	61	
EDP	36	39	21	27	34	41	52	57	59	
Value added as percentage of:										
NDP	100	12.1	12.6	16.7	0.8	7.1	26.2	14.5	9.9	
EDP	100	8.9	8.8	18.7	1.1	9.2	32.5	19.1	13.7	-12
(ii)										
Opening balance sheet:										
Ratio of produced assets/all assets, inclusive of non-produced assets	38	43	13	91	92	95	97	34	29	
Percentage changes between opening and closing balance sheets:										
Net capital formation and volume change due to economic use (ecol. valuation)										
Produced assets	4	2	0	11	2	3	4	10	3	
All assets, incl. non-produced assets	1	-4	-2	10	2	3	4	3	1	
Other volume changes in assets, net										
Produced assets	-2	0	0	-16	0	0	0	-5	-1	
All assets, incl. non-produced assets	0	0	8	-14	0	0	0	-2	-1	
Revaluation and environmental value discrepancies										
Produced assets	14	15	13	13	15	13	12	15	15	
All assets, incl. non-produced assets	20	20	13	14	16	13	13	23	24	
Total changes, net										
Produced assets	15	17	13	8	17	15	16	20	17	
All assets, incl. non-produced assets	21	17	19	10	18	16	16	24	24	
(iii)										
Value added-capital ratios, based on:										
NDP	25	20	73	46	35	21	44	49	6	
EDP	7	4	5	32	31	18	36	15	2	

of economic wealth used in production is even more pronounced: particularly in the mining sector, produced assets are only 13 percent of the total value of assets used by economic activities.

Changes in coverage of assets also affect the change over time of economic wealth between opening and closing balance sheets. In section (ii) of Table 8, the total changes in assets are broken down into net capital accumulation (including ecologically valued volume change of natural assets due to economic use), other volume changes and valuation discrepancies due to market price changes and adjustments of ecological assets to market valuation.

The percentage change attributed to net capital accumulation is higher for produced assets (4 percent) than for all assets (1 percent). For other volume changes, however, this relationship is inverted. In the case of the traditional capital concept, other volume changes—due to earthquake damages—cause a reduction of produced assets (−2 percent), while economic wealth based on the broader concept roughly remains unchanged (0 percent). This inversion is mainly the result of the inclusion of new finds of subsoil assets (27.8, see Table 5) in the latter concept.

Discrepancies in valuation amount to 14 percent for produced assets, reflecting the average annual inflation in the country which was assumed to be 15 percent. In contrast to this, the value of all assets is increased by 20 percent. This is mainly due to the inclusion of the asset of cultivated land with its high price increases (331.0 on an opening stock of 1,366.7, see Table 5).

For the individual sectors a basic pattern can be described: total volume changes, defined as the sum of net capital accumulation and other volume changes, are approximately the same between the traditional capital concept and the broader economic wealth concept. There is a marked difference in the case of agriculture: In this sector, the volume of economic wealth increases by 2 percent (net capital accumulation of 2 percent plus other volume changes of 0 percent) when the narrower concept of produced assets is used, while it decreases by 4 percent (net capital accumulation of −4 percent plus volume changes of 0 percent) when volume changes in all assets are taken into account. The latter reduction is the result of the negative effects of land erosion, depletion and pollution (including acid deposition) on natural resources held by agriculture, forestry and fishing.

The combined changes in value added and economic wealth used in economic activities have considerable effects on the productivity or rentability of capital. The different effects on NDP- and EDP-based value added/capital ratios are presented in the last part (iii) of the table. For the economy as a whole, the ratio between value added and capital based on national accounting (NDP) is 25 percent; this is reduced to 7 percent if based on environmental accounting (EDP). For specific sectors, the differences are even larger. In mining, the value added/capital ratio based on NDP is 73 percent while for EDP it is only 5 percent. For agriculture, the ratio is reduced from 20 percent to 4 percent and for other services from 49 percent to 15 percent. These are significant changes in productivity or rentability indicators which might prompt a reassessment of investment policies as far as capital allocation to economic sectors is concerned. To the extent that environmental costs are also included in (internal) business accounts, new EDP-based measures might also affect micro-economic investment decisions.

References

Ahmad, Y. J., El Serafy, S., and Lutz, E. (eds.). *Environmental Accounting for Sustainable Development*, The World Bank, Washington, D.C., 1989.

Alfsen, K. H., Bye, T., and Lorentsen, L., *Natural Resource Accounting Analysis, The Norwegian Experience, 1978-1986*, Central Bureau of Statistics of Norway, Oslo, 1987.

Ayres, R. U., *Resources, Environment and Economics*, New York, 1978.

Ayres, R. U. and Kneese, A. V., Production, Consumption and Externalities, *American Economic Review*, Vol. 59, No. 3, 282-297, June 1969.

Bartelmus, P., Accounting for Sustainable Development, United Nations, Department of International Economic and Social Affairs, Working Paper No. 8, New York, 1987.

———, *Accounting for Sustainable Growth and Development: Structural Change and Economic Dynamics* (in preparation).

Blades, D. W., Measuring Pollution within the Framework of the National Accounts, in Ahmad, Y. J., El Serafy, S., and Lutz, E., 26-31, 1989.

Daly, H. E., Sustainable Development: From Concepts and Theory towards Operational Principles, *Populations and Development Review* (in preparation).

———, Sustainable Development: From Concepts and Theory towards Operational Principles, (in preparation).

Drechsler, L., Problems of Recording Environmental Phenomena in National Accounting Aggregates, *Review of Income and Wealth*, 22 (3), 239-252, 1976.

El Serafy, S., The Proper Calculation of Income from Depletable Natural Resources, in Ahmad, Y. J., El Serafy, S., and Lutz, E., 10-18, 1989.

Harrison, A., Environmental Issues and the SNA, *Review of Income and Wealth*, 35 (4), 377-388, December 1989.

Hueting, R., *New Scarcity and Economic Growth: More Welfare Through Less Production!* 1980.

Hueting, R. and Leipert, C., Economic Growth, National Income and the Blocked Choices for the Environment, Wissenschaftszentrum Berlin für Sozialforschung (discussion paper), Berlin, 1987.

Institut National de la Statistique et des Etudes Economiques *Les Comptes du Patrimoine Naturel*, Les collections de l'INSEE, Ser. C, 137/138, INSEE, Paris, 1986.

———, *Les Comptes Satellites de l'Environnement, Méthodes et Résultats*, Les collections de l'INSEE, Ser. C 130, INSEE, Paris, 1986a.

Isard, W., Some Notes on the Linkage of the Ecologic and Economic Systems, *Regional Science Association Papers*, 22, 85-96, 1969.

Leipert, C. National Income and Economic Growth: The Conceptual Side of Defensive Expenditures, *Journal of Economic Issues*, 23 (3), 843-856, 1989.

Leontief, W., National Income, Economic Structure and Environmental Externalities, in Moss, M. (ed.), *The Measurement of Economic and Social Performance*, Studies in Income and Wealth, Vol. 38, NBER New York, London, 565-578, 1973.

Lemaire, M. Satellite Accounting: A Solution for Analysis in Social Fields, *Review of Income and Wealth*, 33 (3), 305-325, 1987.

Lutz, E. and El Serafy, S., Recent Developments and Future Work, in Ahmad, Y. J., El Serafy, S., and Lutz, E., 88-91, 1989.

Martinez, A. R., Ion, D. C., De Sorcy, G. J., Dekker, H., and Smith, S., Classification and Nomenclature System for Petroleum and Petroleum Reserves, Twelfth World Petroleum Congress, Houston, 1987.

OECD Environmental Policy Benefits: Monetary Valuation, study prepared by D. W. Pearce and A. Markandya, Paris, 1989.

Olsen, M. The Treatment of Externalities in National Income Statistics, in Wingo, L. and Evans, A. (eds.), *Public Economics and the Quality of Life*, Baltimore, Md., 1977.

Pearce, D. W., Markandya, A., and Barbier, E., *Blueprint for a Green Economy*, London, 1989.

———, *Sustainable Development: Economy and Environment in the Third World*, London, 1990.

Peskin, H. M., A Proposed Environmental Accounts Framework, in Ahmad, Y. J., El Serafy, S., and Lutz, E., 65-78, 1989.

Pezzey, J., Economic Analysis of Sustainable Growth and Sustainable Development, World Bank, Environment Department Working Paper No. 15, Washington, D.C., 1989.

Repetto, R. and others, Wasting Assets: Natural Resources in the National Income Accounts, World Resources Institute, June 1989.

Schäfer, D. and Stahmer, C., Input-output Model for the Analysis of Environmental Protection Activities, *Economic Systems Research*, 1 (2), 203-228, 1989.

———, Conceptual Considerations on Satellite Systems. *Review of Income and Wealth*, 36 (2), 167-176, 1990.

Stahmer, C. Cost-Oriented and Welfare-Oriented Measurement in Environmental Accounting, paper presented at the Fifth Karlsruhe Seminar on Models and Measurement of Welfare and Inequality, August 12-19, 1990.

Teillet, P. A Concept of Satellite Accounts in the Revised System of National Accounts, *Review of Income and Wealth*, 34 (4), 411-439, 1988.

United Nations, *A System of National Accounts*, United Nations publication, No. E.69.XVII.3, 1968.

———, *Draft Guidelines for Statistics on Materials/Energy Balances*, United Nations Publication No. E/CN.3/492, 29 March, 1976.

———, *Provisional International Guidelines on the National and Sectoral Balance-Sheet and Reconciliation Accounts of the System of National Accounts*, United Nations publication, No. E.77.XVII.10, 1977.

———, *The Feasibility of Welfare-Oriented Measures to Supplement the National Accounts and Balances: A Technical Report*, United Nations Publication Series F, No. 22, New York, 1977.

———, *Revised System of National Accounts, Prelimary Draft Chapters, Provisional*, future ST/ESA/STAT/SER.F/2/Rev. 4, February 1990.

———, *International Standard Industrial Classification of all Economic Activities (ISIC)*, United Nations Publication, Series M, No. 4, Rev. 3, New York, 1990a.

———, *Concepts and Methods of Environment Statistics, Statistics of the Natural Environment—A Technical Report*, United Nations publication, in preparation.

———, *Central Product Classification* (CPC), United Nations publication, in preparation.

Uno, K., Economic Growth and Environmental Change in Japan-Net National Welfare and Beyond, in Archibugi, F. and Nijkamp, P. (eds.), *Economy and Ecology: Towards Sustainable Development*, 307-332, Dordrecht, 1989.

Ward, M., *Accounting for the Depletion of Natural Resources in the National Accounts of Developing Countries*, OECD Development Centre Publication, Paris, 1982.

World Commission on Environment and Development, *Our Common Future*, Oxford University Press, Oxford, 1987.

Chapter 4

Development of an algorithm for the compilation of national accounts and related systems of statistics

DEVELOPMENT OF AN ALGORITHM FOR THE COMPILATION OF NATIONAL ACCOUNTS AND RELATED SYSTEMS OF STATISTICS*

BY JAN W. VAN TONGEREN

United Nations Statistical Office

The present article includes a proposal for a national accounts algorithm to be applied in the computerization of the national accounts compilation process. While aiming at the estimation and reconciliation of data from different statistical sources, it is based on the application of a linear programming technique applied to a system of identities and inequalities that define the accounting and analytical relations between the data categories of a national accounting framework. The technique is flexible in the sense that it can be used with any configuration of available statistical sources and data requirements of the national accounts. The algorithm is illustrated graphically with help of a simple example and thereafter applied to an extended but still simple national accounting scheme for Suriname with data for 1965 that was compiled by the author many years ago. As the present study is only a first step in the development of the algorithm, more work is needed to make it operational and the last section of the article includes suggestions about the direction of that further work.

INTRODUCTION

National accounts compilation procedures are increasingly being computerized, either partly or fully. Such computerization requires some kind of formalization of the compilation procedures, so that they can be translated into computer programmes. Without an established theory of estimation in such complex statistical structures as national accounts, there is the distinct possibility that such computer programmes will differ from country to country, reflecting different manual compilation procedures in the past, and depending on the data sources available and data requirements of the particular national accounts system. Thus, experiences of one country may be of very little help to another as the three elements mentioned may differ considerably among countries. Furthermore, as such computerized systems are presently being designed mainly for developed countries, they tend to be very complex given the many sources available and data required, and are therefore not very useful for countries with much less statistical development.

The present study responds to these concerns by developing a flexible compilation algorithm that is based on the application of a linear programming technique to a system of identities and inequalities that define the functional relations between the variables or data categories included in a national accounting framework. The inequalities which correspond to relations among national accounts categories, and between those and the data available from the basic statistical sources, are similar to the ones that constitute the "assumptions" that

*The views expressed in this study are those of the author and do not necessarily coincide with the official views of the United Nations or any of its member governments.

are used at present to compile national accounts. The difference between the present compilation procedures and the algorithm described here is that the assumed relations are defined within intervals. As this would result in interval estimates of national accounts items, the algorithm incorporates the calculation of so-called central values which are similar to the national accounts estimates arrived at through the traditional compilation procedures. The algorithm is defined such that it can be adapted to any configuration of data sources and data requirements. It is a systematic method that can easily be computerized, given the available software packages on linear programming. The method lends itself in particular to national accounts frameworks that are in their earlier stages of development.

Following this introduction, the paper consists of three sections. In Section I the algorithm is described in abstract terms; it starts with the presentation of the present compilation procedures in formal mathematical terms, and derives the proposed or modified algorithm by adjusting the format of the present compilation procedure. Section II applies the algorithm to two examples. The first one is a very simple one, which serves to illustrate the method graphically. The next application is to a more realistic—but still simple—national accounts framework for Suriname related to data for 1965 that was compiled by the author many years ago. The applications show how maximum and minimum as well as central values can be derived for those national accounts categories for which no direct information is available or for which basic statistical sources do not provide adequate information in terms of coverage and definition of the items. The section ends with a sensitivity analysis, which shows how interval estimates and central values are affected if presumed relations involving unknown national accounts categories are replaced in the course of time by estimates based on direct information from statistical sources. Section III discusses conceptual and operational issues which need to be further examined before the proposed algorithm can be effectively applied to actual compilation procedures for national accounts and related statistics.

I. The Formal Structure Of The National Accounts Compilation Procedure

A. *Present National Accounts Compilation Procedure*

The data content of a national accounts scheme and also the national accounts compilation procedures to derive those data are determined by the concepts and classifications used as well as the relations defined between the concepts or variables of the scheme and those available in basic statistical sources. The emphasis of this paper is on the relations between the concepts, as these are the main ingredients of any compilation algorithm. Four kinds of relations can be distinguished:

(i) First there are the identities defined between the variables of the scheme itself. These include the definitions of GDP in terms of expenditure components, activity contributions, income shares and as the difference between gross output and intermediate consumption or as the balance between income and expenditure totals and subtotals in the institutional sector accounts of the national accounts. These identities could be called definitional relationships, as they define one concept or one group of concepts in terms of another.

(ii) A similar group of relationships is defined between information categories in the basic statistical sources and the generally more aggregate categories of the national accounts scheme. These relationships are called here aggregational relations, as they are used in the derivation of the macro-economic categories of the national accounts scheme by aggregating the micro-categories available in the basic sources. For example gross output, intermediate consumption and value added components of the mining sector may be derived—if information is available—by simple aggregation of gross output and the cost components of all establishments that could be classed as mining establishments. Another example is the taxation categories that are used in government statistics, which are derived by simple agregation of the appropriate very detailed tax items that can be identified in the government administrative records.

(iii) A third group of what will be called analytical relations between the variables of the national accounts scheme are those that reflect institutional arrangements such as taxation functions, behaviouristic patterns such as consumption, investment, saving and import functions, and technical relations that are defined between the goods and services flows of a national accounting scheme and in particular between gross output, cost components and capital formation.

(iv) The fourth type are those relations that are typical for the present national accounts compilation process. They are the ones that define the adjustments to the data of the basic sources in order to arrive at macro-economic concepts of the national accounts. They are used in order to correct for lack of coverage or detail in the basic source information, and they involve marginal adjustments needed in order to adapt statistical source concepts to the different coverage defined in the national accounts. These relations are similar to the analytical relations. In fact, use is often made of the analytical relationships in order to correct the deficiencies in the basic data. For example, employment data are used to inflate the survey results for a limited number of production establishments to arrive at the universe as a whole, or production data of similar establishments are utilized to obtain further detail on the intermediate consumption of those establishments for which such information was lacking. Both types of adjustments are based on technical relations between gross output, intermediate consumption and employment data.

If data needed for the derivation of the national accounts aggregates were available in the basic sources, the only type of relation which would be used, and thus would define the national accounts compilation procedure, would be the aggregational relations between the data of the basic sources and the macro concepts. There would be no need to take the national accounts definitional relationships or identities explicitly into account as these would already be present in the basic sources in a disaggregated form. Simple aggregation would guarantee that such identities would then automatically hold in the macro-economic framework of the national accounts as well. Also there would be no need to account explicitly for the analytical relations as they would be available in the basic sources in their micro-format and through simple aggregation would be reflected in the macro concepts. It is also obvious that the adjustments to the basic data would not be relevant, as in the ideal situation, concepts and classifications of the basic sources would coïncide with those of the macro framework of the accounts or at least provide the basic elements for deriving the macro-economic concepts.

In practice, the ideal situation does not exist. Basic sources are lacking in coverage and in detail, the concepts do not coincide with those of the national accounts, and there are inconsistencies between data of different sources. This implies that the macro-economic definitional and analytical relationships will not be fully present in the basic sources and therefore need to be imposed by incorporating those relations in the compilation process. Similarly, adjustments to the basic data are needed in order to bring them in line with the macro-economic concepts.

The present national accounts compilation process can be defined symbolically as the solution to the following set of equations:

$$A1 \cdot X = 0$$

$$X = A2 \cdot Y$$

$$B1_p \cdot X = 0$$

$$X = B2_p \cdot Y.$$

The symbols X and Y are vectors, respectively, of variables in the macro framework of the national accounts and of variables available in the basic sources. Vector Y includes many more components than X, as it reflects all the detail of information in the basic sources. $A1$, $A2$, $B1$, $B2$ are matrices of coefficients. Matrix $A1$ reflects the definitional relationships among the variables of the national accounting framework. It includes only three types of values, i.e. $+1$, -1 and 0. The same holds for the coefficients included in $A2$ of all aggregational relationships between the variables of the basic sources and those of the macro framework that can be defined with help of available basic information. $B1$ includes the coefficients of the analytical relationships that hold between the data within the national accounts framework. The coefficients included can have other values than the ones mentioned for matrices $A1$ and $A2$. Matrix $B2$ is the matrix which includes the coefficients that define the adjustments in terms of coverage, detail and concepts that have to be made to the basic source data categories in order to obtain the categories that are required in the macro framework of the accounts. This matrix may also include coefficients other than the ones mentioned for the definitional and aggregational matrices $A1$ and $A2$. Present national accounts practices generally use ratios between pairs of variables in defining the coefficient matrices $B1$ and $B2$. This implies that such matrices are of a very simple type, which only include two coefficients per row, one of which is equal to 1. The ratios are between data available in the basic sources and unknown categories of the macro framework and are therefore mainly of the adjustment type $B2$. Those analytical relations between pairs of variables within the national accounts framework (matrix $B1$), where one of the variables is derived from the basic data through simple aggregation ($A2$) may also be included. Present national accounts practices do not involve other analytical relations between unknown variables in the national accounts framework.

The definitional relations defined by matrix $A1$ will include only those identities that are not present in the basic statistical sources, either because of lack of detail or coverage or conceptual discrepancies between the basic source categories and those of the macro-economic national accounting framework, or because of inconsistencies between the data of different statistical sources. Identities that are already present in the basic sources do not need to be repeated in defining the compilation procedure. In principle, all independent aggregational relationships between basic sources and the macro framework are included in $A2$. When adding up the number of relations thus incorporated in defining the compilation procedure, one generally would arrive at a number that is smaller than the number of national accounts variables. In order to obtain a "solution" of the national accounts model, the number of relations included is then increased by a selection from the available analytical and adjustment relations, so that the total number of functional relationships is equal to the number of national accounts variables to be compiled. This partial selection is reflected in the p subscripts of $B1$ and $B2$. The national accounts "solution" arrived at is therefore not a unique solution, but changes with the selection of relationships included in $B1$ and $B2$. Present compilation procedures try to offset this disadvantage by checking, in an *ad hoc* manner, estimates based on one selection of analytical and adjustment functions with other relations that also should hold. For instance, if the use of the commodity flow method results is unacceptable values for value added in particular industries, the assumptions built into the commodity flow method are reexamined and revised so that a more "acceptable" level of value added is obtained for those sectors. This iterative checking of estimates is, however, not necessarily a systematic procedure which would involve all relations that could potentially be included in $B1$ and $B2$. The result is that national accounts data may show up relationships that are not acceptable when used in particular types of analysis. If the national accountant had incorporated those relationships in his compilation procedure, such deficiencies might have been avoided.

B. *General Principles of the Modified Compilation Algorithm*

The major restriction of the present compilation procedure used in national accounting is the one that determines the number of analytical and adjustment relations that can be incorporated in a systematic manner. It is proposed in this paper that this restriction be removed by replacing the single values of the analytical and adjustment parameters in the $B1$ and $B2$ matrices by intervals that are defined by maximum and minimum values of those parameters. This implies that the $B1$ and $B2$ matrices would be replaced by pairs of matrices, one including the minimum values of the $B1$ and $B2$ coefficients and another the maximum values of those parameters. The equality signs in the two expressions that include $B1$ and $B2$ will then be replaced by inequality signs, such that the expressions including the minimum values of the two matrices would be presented with $\geq$ signs and the maximum values with $\leq$ signs. The modified compilation procedure could then be defined as finding the maxima and minima of the variables included in the X-vector which are within the bounds of the following set of identities and inequalities.

$$\begin{aligned} A1 \cdot X &= 0 \\ X &= A2 \cdot Y \\ B1\,\mathrm{min} \cdot X &\geq 0 \\ B1\,\mathrm{max} \cdot X &\leq 0 \\ X &\geq B2\,\mathrm{min} \cdot Y \\ X &\leq B2\,\mathrm{max} \cdot Y. \end{aligned}$$

In the symbolic presentation above, $B1$ and $B2$ in their maximum as well as in their minimum value versions no longer include the restrictive subscript p. This means that in principle the new procedure may incorporate as many analytical and adjustment relations as required or available, provided of course that the intervals do not conflict with each other. The ultimate intervals of national accounts estimates that result from this computation procedure will then be consistent with many more analytical and adjustment relations than the present procedure could ever satisfy. This formulation furthermore opens the possibility that not only will relations be included between known basic source variables and unknown national accounts aggregates and between unknown national accounts aggregates and aggregates that are derived with the help of the simple aggregational relationships as mentioned above, but also that analytical relations can be defined among two or more unknown variables in the national accounts scheme itself. This enhances the analytical usefulness of the national accounts variables, as they will be consistent with many more analytical and adjustment relationships that have never entered the national accounts compilation process before. There is furthermore the possibility of incorporating, in addition to ratios that are presently used in national accounts compilation procedures, analytical functions that have been derived as a result of econometric or related research. The latter possibility would contribute to further integration of national accounts compilation with econometric and related research.

The application of the linear programming algorithm determines a multi-dimensional region within which all feasible X-solution vectors are located. The minimum and maximum values of the components of X are to be found on the edges of this region, and are included in what might be called extreme value vectors. In order to arrive at singular values for each of the components of X similar to the estimates that result from the present national accounts compilation procedures, the algorithm should include additionally the derivation of a solution vector of so-called central values which might be defined as a function of the extreme value vectors that include the minima and maxima of the components X. The precise definition of this functional relationship depends on the criteria that single valued estimates should satisfy. No definite conclusions regarding these criteria have been reached in this study, but preliminary suggestions on the theoretical concept of central values are presented in section III, and a practical illustration is worked out in the examples of section II.

II. Illustration Of The Use Of The Modified Algorithm In The Compilation Of National Accounts Data Schemes

A. *Simple Example Illustrating the Present and Modified Compilation Procedures*

The more abstract formulations in the previous section of the present and modified national accounts compilation procedures will be first illustrated with the help of a very simple example. The present compilation is reflected in Table 1 and the modified algorithm in Table 2. Both tables assume a definitional relationship based on the expenditure approach to GDP. The variables indicated with a bar in the two tables (for instance $\bar{C}_g$) are those national accounts aggregates that are derived directly from available basic statistical surveys (e.g. economic censuses and surveys, government records and foreign trade and balance of payments information) with help of the simple aggregational relationships earlier

TABLE 1

Alternative Solutions to a Simple National Accounts Model based on Present National Accounts Compilation Practice

Definitional relationship:

$$\bar{Y} = Ch + \bar{C}g + If + \bar{I}s + \bar{E}x - \bar{M}$$

$\bar{Y}$ = GDP at market prices (230)
Ch = Private final consumption expenditure
$\bar{C}g$ = Governmental final consumption expenditure (37)
If = Gross fixed capital formation
$\bar{I}s$ = Increases in stocks (6)
$\bar{E}x$ = Exports of goods and services (120)
$\bar{M}$ = Imports of goods and services (200)

or

$$Ch + If = 274$$

Analytical and adjustment relations:

$$Ch = 3\ If$$
$$If = 0.2\ \bar{Y}$$
$$Ch = 1.05\ \bar{C}ho\ (\bar{C}ho = 210)$$

Alternative solutions of Ch and If:

	$Ch = 3\ If$	$If = 0.2\ \bar{Y}$	$Ch = 1.05\ \bar{C}ho$
Ch	205.5	228.0	220.5
If	68.5	46.0	53.5

TABLE 2

MAXIMUM, MINIMUM AND CENTRAL VALUES DERIVED FROM A SIMPLE NATIONAL ACCOUNTS MODEL ON THE BASIS OF THE MODIFIED NATIONAL ACCOUNTS COMPILATION ALGORITHM

Definitional, analytical and adjustments relations:

$$Ch + If = 274$$
$$2.5 \le Ch/If \le 8$$
$$0.1 \le If/\bar{Y} \le 0.3$$
$$1.03 \le Ch/\bar{C}ho \le 1.08$$

or

$$Ch + If = 274$$
$$Ch - 2.5\ If \ge 0$$
$$Ch - 8\ 8\ If \le 0$$
$$If \ge 0.1\ \bar{Y}$$
$$If \le 0.3\ \bar{Y}$$
$$Ch \ge 1.03\ \bar{C}ho$$
$$Ch \le 1.08\ \bar{C}ho$$

Maximum, minimum and central values:

	Minimum value			Maximum value	
	Ex ante	*Ex post*	Central value	*Ex post*	*Ex ante*
Ch	—	216.3	221.6	226.8	—
If	—	47.2	52.4	57.7	—
Ch/If	2.50	3.75	4.23	4.81	8.00
$If/\bar{Y}$	0.10	0.21	0.23	0.25	0.30
$Ch/\bar{C}ho$	1.03	1.03	1.06	1.08	1.08

mentioned. Total GDP (Y) in the example is measured with help of the alternative production or income approaches. The aggregational relationships have not been explicitly incorporated in this example; instead the resulting values of the barred variables have been included and indicated in brackets ( ) following the definitions of those variables. The unknown variables in this national accounts model are gross fixed capital formation (*If*) and private final consumption expenditure (*Ch*). Given the values of the known variables, the definitional relationship reduces to a simple sum of *If* and *Ch* equal to 274.

Three types of analytical and adjustment relations are used to estimate the two unknown variables. The first one, which is based on the commodity flow method applied to further detail of gross output, imports and known expenditure categories, results in a distribution of the commodity flows between gross fixed capital formation (*If*) and private final consumption expenditure (*Ch*). The second one relates the value of capital formation to GDP on the basis of past measurements of capital output ratios. A third relation assumes a growth rate of private final consumption expenditure (*Ch*) as compared with its last year's value and based on past experience with regard to the link between the growth rates of consumption and GDP.

In the application of the present compilation methodology in Table 1 the parameters of the analytical and adjustment relations have been given specific values. It has been assumed that the application of the commodity flow method results in a value of private final consumption (*Ch*) that is 3 times the amount of gross fixed capital formation (*If*). Past experience has furthermore shown that gross fixed capital formation (*If*) is 20 percent of GDP ($\bar{Y}$) and that the growth rate of consumption might be 5 percent since last year. As only one of these three types of information is needed jointly with the national accounts identity in order to solve for the value of *If* and *Ch*, there are in fact three solutions to the model and all are different. The commodity flow assumption leads to the lowest value of *Ch* and the highest value of *If*, while the assumption regarding the capital-output ratio results in the highest value of *Ch* and the lowest value of *If*. The assumption about the growth of consumption gives results that are intermediate between the two levels.

It is obviously unacceptable that the three assumptions would lead to such different results. This reflects on the quality of the assumptions as well as on the data regarding the known variables. National accountants using the present methodology are aware of this problem and to a limited extent try to resolve it by confronting the varying solutions and making adjustments to the assumed parameters and data in an iterative manner, so that ultimately one acceptable solution evolves. If no reconciliation is feasible, they would generally accept one of the solutions as the most reliable one with which other data are reconciled, or introduce statistical discrepancies to reflect the different values of the solutions.

The modified algorithm as illustrated in Table 2 avoids the irreconciled or irreconcilable solutions by replacing the fixed values of the parameters by ranges and the equality signs in the analytical and adjustments relations by inequality signs. Instead of assuming that consumption is 3 times capital formation it is defined within a range of 2.5 to 8 times the values of capital formation, taking into account the reliability of the data and assumptions used in the commodity flow method. Similarly, the ratio between capital formation and output is not fixed at 0.2 but defined between 0.1 and 0.3, and the growth rate of consumption (based on a consumption figure in the base period $\bar{Cho} = 210$) is assumed to be between 3 and 8 per cent, reflecting variations in the past in these parameters. The solution to the modified compilation model presented in table 2 is not a unique one, but rather consists of a set of intervals for the variables and parameters of the model. These ranges which are called *ex post* intervals are distinguished from the *ex ante* intervals of the parameters which are defined at the outset of the model. The table also presents singular values for each of the variables similar to those resulting from the traditional compilation methods. These are called central values and are defined as the components of a solution vector that is the unweighted average of all extreme solution vectors that include *ex post* minima and maxima of the unknown variables and parameters in conformity with the restrictions of the compilation model.

As the example is simple, the *ex ante* and *ex post* intervals and the central values resulting from the application of the modified compilation algorithm in Table 2 can be graphically illustrated. This is done in diagram A. The co-ordinates of the diagram correspond to the two unknown variables in the compilation model, i.e. *Ch* is presented on the horizontal axis and *If* vertically. The GDP identity is represented by one line in the diagram and the *ex ante* analytical and adjustment restrictions by areas between pairs of lines that are defined by pairs of inequalities of the model. The solution is to be found on that part of the line of the definitional identity of GDP that coincides with an area commonly defined

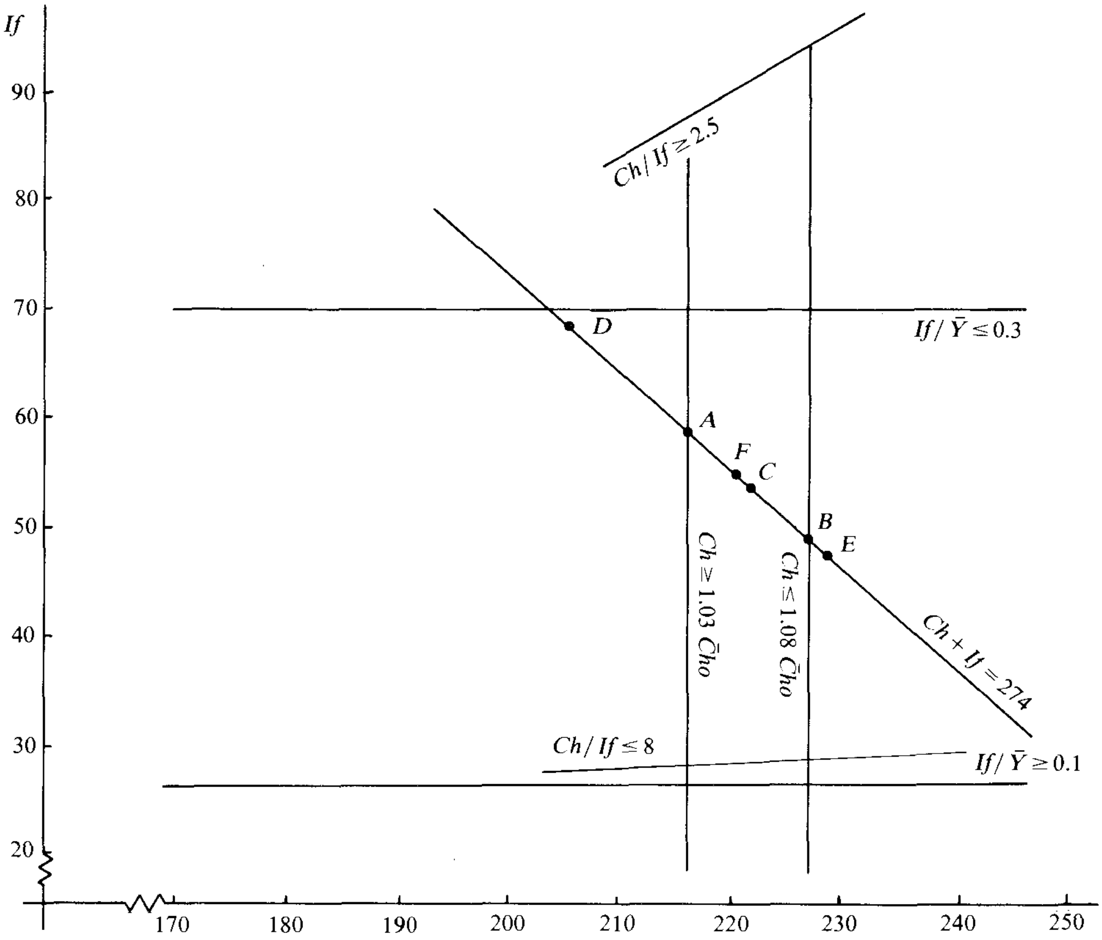


Diagram A. Graphical Representation of the Simple National Accounts Compilation Model

by all analytical and adjustment relations, i.e. between points A and B. Point C is the central value which is defined as the unweighted average of solution vectors A and B. Points D, E and F represent the three solutions shown in Table 1, that would be the result of the present compilation procedure. Each of these three solutions is also on the GDP identity line. Point A corresponds to the solution vector (*Ch*, *If* = 216.3, 57.7), B to the solution vector (226.8, 47.2) and the central value C vector is (221.6, 52.4). The alternative solution vectors of the present compilation method D, E and F are respectively (205.5, 68.5), (228.0, 46.0) and (220.5, 53.5).

Table 2 and diagram A show that there is a considerable reduction in the intervals of two of the three parameters after confrontation of the *ex ante* intervals of the analytical and adjustment relations and the definitional identity. The 2.5-8 *ex ante* interval of Ch/If reduces *ex post* to 3.75-4.81 and the 0.10-0.30 interval of $If/\bar{Y}$ reduces to 0.21-0.25. The $Ch/\bar{C}ho$ *ex post* interval remains the same compared to the *ex ante* interval and thus determines the *ex post* intervals of all other parameters and variables in the model. The reason that there are two of such limiting *ex ante* extremes—i.e. the minimum and maximum of $Ch/\bar{C}ho$—is because there are two variables (*If* and *Ch*) to be solved. The central values are simple averages of the two solution vectors A and B, and therefore automatically satisfy the definitional relationships as well as the *ex post* ranges of variables and parameters. From the presentation of the diagram it also becomes clear that two of three alternative solutions D and E that result from the traditional compilation procedure are located outside the *ex post* range AB. These solutions therefore do not satisfy simultaneously all *ex ante* ranges of the parameter intervals, which illustrates the possibility of arriving at erroneous results when applying the traditional compilation methods.

B. *Application of the Modified Algorithm to an Extended National Accounts Scheme*

The simple example presented above was sufficient to illustrate the general principles of the modified algorithm. However, it cannot reflect the implications for the solution procedure when applied to a more realistic national accounting scheme which would involve many more variables and parameters. To illustrate this, use is made of preliminary data that were compiled by the author for Suriname covering the year 1965, in an earlier attempt to develop the compilation algorithm presented in this study. The data included were later on used and further improved in the actual compilation of national accounts for Suriname for that year on the basis of the traditional national accounts compilation procedure.

The national accounts scheme is limited to production account information. The data categories directly estimated and used in the compilation are presented in sections (a), (b), (c) and (d) of Table 3 below. Section (a) includes production accounts for four key sectors: agriculture, bauxite-aluminum industry, other industries and the public sector. Agriculture includes the traditional small scale agriculture, as well as the larger plantations that apply more advanced production methods. The bauxite-aluminum industry consists of a limited number of enterprises which dominate the mineral exploitation of bauxite and its processing to alumina and aluminum and are also a major source of supply of electricity for the country. Other industries cover the remaining private enterprises that operate in such diverse production processes as timber production and processing, construction, simple production processes as bottling of imported beverages, and also all service activities which are mainly operated on a small scale. The public sector includes general government activities as well as government enterprises such as the postal and telephone system, port authority, Suriname Airlines, agricultural experimental stations, railroad and bus services, government car repair shops, etc. The production accounts only include aggregate categories of output and cost: gross output, intermediate consumption, and value added broken down by compensation of employees and the rest of value added covering operating surplus as well as depreciation, indirect taxes and subsidies. In each of the production accounts imported goods and services are separately identified within intermediate consumption.

Section (b) of Table 3 presents the breakdown of GDP by cost components. It includes compensation of employees for all sectors together and shows a breakdown of the remaining value added by gross operating surplus and indirect taxes and import duties. Section (c) of the table includes a breakdown of GDP

TABLE 3

DATA CATEGORIES OF THE EXTENDED NATIONAL ACOUNTING SCHEME

	Agriculture	Bauxite-Aluminum Sector	Other Industries	Public Sector
(a) PRODUCTION ACCOUNTS				
Gross output	48,628	172,311	Xo	71,634
Intermediate consumption	9,505	103,472	Co	30,182 ($\bar{C}g$)
Of which: Imports	4,283 ($\bar{M}ca$)	77,838 ($\bar{M}cb$)	Mco	
Value added	39,163 $\bar{Y}a$)	68,839	Yo	41,452
Compensation of employees	Sa	25,902 ($\bar{S}b$)	So	41,452 ($\bar{S}g$)
Other value added	Pa	42,937 ($\bar{P}b$)	Po	—
(b) COST COMPONENTS OF GDP				
Compensation of employees			S	
Other value added			P'	
Indirect taxes			37,471	
Of which: Import duties			28,380	
Operating surplus, gross			P	
Gross domestic product			Y	
(c) GDP BY TYPE OF FINAL EXPENDITURES				
Final consumption expenditure				
Households			Ch	
Government			59,734 ($\bar{X}g$)	
Gross output			71,634	
Minus: Government sales			−11,900	
Of which: Imports			43,800 ($\bar{M}c$)	
Gross fixed capital formation			If	
Dwellings			9,782 ($\bar{I}fh$)	
Bauxite-aluminum sector			78,061 ($\bar{I}fb$)	
Agriculture and other industries			$Ifao$	
Government			25,703 ($\bar{I}fg$)	
Of which: Imports			Mif	
Changes in stocks			6,014 ($\bar{I}s$)	
Exports			123,049 ($\bar{E}x$)	
Agriculture			11,981	
Bauxite-aluminum sector			87,729	
Other industries			21,882	
Government			1,457	
Minus: Imports			199,100 ($\bar{M}$)	
Gross Domestic Product			Y	
(d) OTHER BASIC DATA				
Compensation of employees of large enterprises in other industries			39,221 ($\bar{S}o$)	
Investments made by large enterprises, other industries			11,848 ($\bar{I}fo$)	
Imports of investment goods identified, other industries			41,600 ($\bar{M}if$)	
Corporate income tax (other than paid by the bauxite-aluminum sector)			1,568 ($\bar{T}dao$)	
Income tax paid by individuals			6,451 ($\bar{T}dh$)	

by categories of final demand, i.e. final expenditure of households and government, gross fixed capital formation, changes in stocks, exports and imports. The import components of final consumption expenditure and gross fixed capital formation are separately identified. Government final consumption expenditures are deconsolidated into gross output minus sales of the public sector (including sales by government enterprises incorporated in the public sector), and gross fixed capital formation and exports are broken down by sector.

The intermediate data available for the exercise was the following. Annual agricultural surveys provided information on gross output by type of agricultural products and a survey of agricultural holdings included data on intermediate consumption. Direct production account data were available from the few enterprises operating in the bauxite–aluminum industry. This included data on gross output, intermediate consumption, imports, the breakdown of value added, exports, gross fixed capital formation and changes in stocks. Government administrative records were the main source of data for the public sector. For the remaining industries an annual survey of large enterprises was conducted in co-operation with an auditing firm, which supplied data on compensation of employees and capital formation only. Foreign trade statistics included all needed detail on exports by each of the sectors distinguished. Among imports of goods it was possible to identify imports for the agricultural and bauxite–aluminum sectors, and a limited number of other items could be identified as either consumption goods or goods destined for capital formation. The remaining import goods could not be further identified. Balance of payments statistics compiled by the Central Bank could distinguish between imports of services for each of the sectors. Additional information available from tax records supplied data on direct taxes paid by individuals, enterprises of the bauxite and aluminum industry, and enterprises operating in agriculture and other industries, as well as total indirect taxes and import duties collected by the government. All direct information on national accounts categories that was used in the application of the algorithm is presented in quantitative form in sections (a), (b) and (c) of Table 3. Data that does not conform to the definitions and/or coverage of the categories of the national accounting scheme but is used in the compilation is described and quantified in section (d) of the table.

To estimate the unknown variables, definitional, analytical and adjustment relations have been defined between known and unknown variables as well as among unknown variables. The functional relationships and the intervals for the parameters are presented in Table 4. The variables that are used in the relations are defined in the four sections of Table 3. The main variables are gross output (X), value added and GDP (Y), intermediate and final consumption (C), gross capital formation (I), exports (Ex), imports (M), compensation of employees (S) and other value added (P). Subscripts and primes (′) are added to denote further subdivisions, sectors and different definitions of the same concept. Known categories of information that could be directly derived from the basic data are indicated with a bar (e.g. $\bar{C}g$) and unknown variables are presented without a bar (e.g. S).

Four definitional relationships have been defined that are all independent of each other and include at least some unknown variables. The first one is the

TABLE 4

IDENTITIES AND ADJUSTMENT AND ANALYTICAL RELATIONS DEFINED IN THE EXTENDED NATIONAL ACCOUNTS SCHEME*

Identities		
	$Pa+\bar{P}b+(Xo+Co-So)+S=Ch+\bar{X}g+(Ifao+\bar{I}fh+\bar{I}fb+\bar{I}fg)+\bar{I}s+\bar{E}x-(\bar{M}g+\bar{M}s)$	(1)
	$\bar{M}g=\bar{M}c+\bar{M}ca+\bar{M}ob+Mco+Mif$	(2)
	$S=Sa+\bar{S}b+So+\bar{S}g$	(3)
	$\bar{Y}a=Sa+Pa$	(4)
Adjustment relations		
	$1 \leq Mif/\bar{M}if \leq 2$ (or *lmif*)	(5)
	$1.5 \leq Ifao/\bar{I}fo \leq 2.5$ *lifao*)	(6)
	$1.9 \leq So/\bar{S}o \leq 2.4$ (or *lso*)	(7)
Analytical relations		
	$0.25 \leq \bar{M}c/(Ch+\bar{C}g) \leq 0.6$ (or *mc*)	(8)
	$0.25 \leq Mif/(Ifao+\bar{I}fh+\bar{I}fb+\bar{I}fg) \leq 0.6$	(9)
	$0.25 \leq Mco/Co \leq 0.6$ (or *mco*)	(10)
	$0.3 \leq Co/Xo \leq 0.7$ (or *co*)	(11)
	$0.2 \leq So/Xo \leq 0.7$ (or *so*)	(12)
	$0.55 \leq Sa/\bar{Y}a \leq 0.95$ (or *sa*)	(13)
	$0.7 \leq Ch/S \leq 0.95$ (or *ch*)	(14)
	$0.8 \leq Ifao/(Xo-Co-So+Pa)$ (or *ifao*)	(15)
	$0.03 \leq \bar{T}dh/S \leq 0.06$ (or *tdh*)	(16)
	$0.03 \leq \bar{T}dao/(Pa+Xo-Co-So) \leq 0.25$ (or *tdao*)	(17)

*For an explanation of the symbols of the variables, refer to Table 3.

identity between expenditure minus imports and the total of income shares of GDP. The second identity shows how the total imports of goods are broken down by goods imported for purposes of intermediate consumption for the three production sectors, imported consumption goods for public and private consumption, and imported capital goods. The third identity stands for the breakdown of total compensation of employees by industrial sectors. The last identity used defines value added of the agricultural sector as the sum of compensation of employees and the non-labour income component of value added in that sector.

Three adjustment relations are defined between the incomplete data of the basic sources and the complete coverage of the items that should be reflected in the national accounts scheme. The first one assumes that actual imports of capital goods are between 1 and 2 times the amount that could be identified as capital goods imports in the foreign trade statistics. The second relation assumes that gross fixed capital formation in agriculture and other industries is between 1.5 and 2.5 times the amount of capital goods purchased by the large industrial enterprises surveyed in co-operation with an auditing firm. It was further assumed that compensation of employees paid out by the large enterprises surveyed should be multiplied by a factor between 1.9 and 2.4 in order to arrive at compensation of employees for the total of other industries. The intervals are either based on expert judgement (using large intervals) or past experience.

Expert judgement and past experience were also used to define lower and upper bounds for ratios reflecting six analytical relations between variables. Identified imports of consumption goods were assumed to be between 0.25 and 0.6 of final consumption by households and intermediate consumption by the government sector. The same interval was assumed for the ratio between total imports of capital goods and capital formation in agriculture and other industries and between imports of goods and services for intermediate consumption by other industries and intermediate consumption of those industries. Intermediate consumption of other industries was assumed to be between 0.3 and 0.7 of gross output of other industries and compensation of employees between 0.2 and 0.7 of gross output of other industries. In agriculture it was assumed that compensation of employees was between 0.55 and 0.95 of total value added of agriculture. Household consumption was considered to be between 0.7 and 0.95 of compensation of employees. Gross fixed capital formation in agriculture and other industries was assumed to be larger than 0.8 of non-labour income generated in these sectors. Direct taxes paid by individuals were fixed between 0.03 and 0.06 of compensation of employees and direct taxes paid by corporations and other enterprises operating in agriculture and other industries between 0.03 and 0.25 of non-labour income generated by those sectors.

The identities and lower and upper bounds of parameters of the adjustment and analytical relations are rewritten in Table 5 in linear programming (lp) format in order to apply a linear programming technique for deriving the *ex post* minima and maxima of the unknown variables of the national accounts model. This technique replaces the graphical analysis earlier applied to the simple model in section II A, which is not suitable for complex national accounts models with a large number of variables. The relations are divided in the table in three groups: the first group are the identities; the second group are the inequalities that include more than one unknown variable; the third group includes those inequalities with one unknown variable which is defined in terms of known estimates directly derived from available statistical sources. The reference codes of the relations used in this table are the same as those of Table 4. The analytical and adjustment relations are presented in pairs of two, which define the minimum ($\geq$) and maximum ($\leq$) values of the unknown national accounts categories in terms of the known ones. The unknown categories are presented on the left-hand side of the equality and inequality signs and the known categories on the right-hand side.

To arrive at the *ex post* minima and maxima, the inequalities of Table 5 are converted to equalities by introducing slack variables that are defined to be equal

TABLE 5

LINEAR PROGRAMMING FORMAT OF THE IDENTITIES AND ANALYTICAL AND ADJUSTMENT RELATIONS OF THE EXTENDED NATIONAL ACCOUNTS SCHEME*

Unknown national accounts variables											
Xo	*Co*	*So*	*Ifao*	*Pa*	*Sa*	*Mif*	*Mco*	*Ch*	*S*		
+1	−1	−1	−1	+1				−1	+1	$= -\bar{Pb} + (\bar{Ifh} + \bar{Ifb} + \bar{Ifg})$	
										$+ \bar{Is} + \bar{Ex} - (\bar{Mg} + \bar{Ms})$	(1)
						+1	+1			$= Mg - (\bar{Mc} + \bar{Mca} + \bar{Mcb})$	(2)
		−1			−1				+1	$= \bar{Sb} + \bar{Sg}$	(3)
				+1	+1					$= \bar{Ya}$	(4)
			−0.25			+1				$\geq 0.25\,(\bar{Ifh} + \bar{Ifb} + \bar{Ifg})$	(9)
			−0.6			+1				$\leq 0.6\,(\bar{Ifh} + \bar{Ifb} + \bar{Ifg})$	
								+1	−0.7	≥ 0	(14)
								+1	−0.95	≤ 0	
−0.2		+1								≥ 0	(12)
−0.7		+1								≤ 0	
−0.8	+0.8	+0.8	+1	−0.8						≥ 0	(15)
−0.3	+1									≥ 0	(11)
−0.7	+1										
	−0.25						+1			≥ 0	(10)
	−0.6						+1			≤ 0	
+1	−1	−1		+1						$\geq \bar{Tdao}/0.25$	(17)
+1	−1	−1		+1						$\leq \bar{Tdao}/0.03$	
						+1				$\geq \bar{Mif}$	(5)
						+1				$\leq 2\,\bar{Mif}$	
			+1							$\geq 1.5\,\bar{Ifo}$	(6)
			+1							$\leq 2.5\,\bar{Ifo}$	
		+1								$\geq 1.9\,\bar{So}$	(7)
		+1								$\leq 2.4\,\bar{So}$	
								+1		$\geq \bar{Mc}/0.6 - \bar{Cg}$	(8)
								+1		$\leq \bar{Mc}/0.25 - \bar{Cg}$	
					+1					$\geq 0.55\,\bar{Ya}$	(13)
					+1					$\leq 0.95\,\bar{Ya}$	
									+1	$\geq \bar{Tdh}/0.06$	(16)
									+1	$\leq \bar{Tdh}/0.03$	

*For explanation of the symbols of the variables and parameters, refer to Tables 3 and 4, respectively.

to the difference between the left and right hand side values of the inequality expressions. The *ex post* extremes are then obtained as the optimum values of objective functions that are subsequently defined in terms of each of the unknown variables separately, and subject to all restrictions defined in the table. The subsequent solutions can be represented by vectors, each of which includes the maximum or minimum of at least the one variable that was optimized and extreme or non-extreme values for all other unknown variables. These so-called extreme value vectors are then used to derive a central value vector which contains for each of the unknown variables the unweighted averages of the corresponding values included in each of the extreme value vectors. The values of the components of the central value vector thus defined automatically satisfy the *ex ante* restrictions of the national accounts model and are also consistent with the *ex post* maxima and minima of the variables.

The complexity of the modified algorithm is determined by the size of the lp matrix. This size is dependent on the number of unknown variables and slack variables and on the number of identities and inequalities which define the compilation procedure. It should be noted here that the inequalities always appear in pairs, which considerably increases the size of the matrix. Not reflected in the lp matrix are the inequalities with only one unknown variable, as these can be redefined as minimum and maximum values of unknown variables, and accommodated without enlarging the lp matrix by using a modified lp technique [1] based on upper and lower bounds of unknown variables. Further reductions in the size of the matrix were effected by not including those unknown variables that can be defined as linear combinations of the other variables. This reduced the number of unknown variables from the 16 explicitly presented in Table 3 to the 10 included in the lp presentation of Table 5. For instance, GDP (Y) was not included by defining it as the sum of the value added components in Tables 3(a) and 3(b) or as the sum of the expenditure categories presented in Table 3(c). Total operating surplus (P) was left out by defining it as the sum of operating surplus of each of the sectors separately. Other value added of other industries (P_0) was not included by defining it as the difference between gross output and intermediate consumption and compensation of employees of other industries.

The results of the lp iteration have been presented in Table 6. The first section of the table shows the *ex post* extremes as well as central values of the unknown national accounts categories. Also included in this part of the table is the relative size of the *ex post* intervals which, for each of the variables is expressed as a percentage of their central values. The second part of the table shows the central values and *ex ante* and *ex post* minima and maxima of the parameters or ratios defined in Table 4, and also includes in the last two columns the relative sizes of the *ex ante* and *ex post* intervals expressed as percentages of the central values of these parameters.

Conclusions similar to those for the simple example examined in section IIA can be drawn from Table 6 for the application of the lp algorithm to the extended national accounts model. Of the *ex ante* extremes of the parameters, 10 remain unchanged when measured *ex post* (i.e. the minimum and maximum values of *lifao*, *lso* and *sa* and the minimum values of *mco*, *co*, *ch* and *ifao*). The number of these extremes correspond to the number of unknown variables in the national accounts model. They are the ones that determine changes in the *ex post* as compared to the *ex ante* values of the remaining extremes of the parameters shown in the second section of table, and also the *ex post* intervals of the national accounts variables presented in the first section. Some of the parameter intervals have been considerably reduced as compared to their *ex ante* length. This applies in particular to the parameters of imports (*lmif*, *mif*, *mc*, *mco*), final consumption (*ch*), intermediate consumption of other industries (*co*) and the tax parameters (*tdh* and *tdao*). With regard to the unknown variables in the model (the first section of Table 6), the relative *ex post* intervals are fairly small for imports (*Mco*, *Mif*), gross output, intermediate consumption and value added of other industries (*Xo*, *Co*, *So*), for the main expenditure and cost components of GDP (*Ch*, *If*, *S*, *P*) and for GDP (Y) itself. Further breakdowns of items by detailed cost components and expenditure categories (*Sa*, *Pa*, *So*, *Po* and *Ifao*) have *ex post* intervals that are much wider. The central values presented in the table are different (sometimes lower, sometimes higher) from the simple mid-points

TABLE 6

Ex Ante AND *Ex Post* MINIMUM, MAXIMUM AND CENTRAL VALUES OF VARIABLES AND PARAMETERS OF THE EXTENDED NATIONAL ACCOUNTS SCHEME, PRIOR TO AND OBTAINED AS A RESULT OF THE APLICATION OF THE MODIFIED NATIONAL ACCOUNTS COMPILATION ALGORITHM

	Minimum Values		Central Values	Maximum Values		Interval Length as a Percentage of the Central Value	
	Ex ante	*Ex post*		*Ex post*	*Ex ante*	*Ex ante*	*Ex post*
Variables							
Sa		21,540	32,914	37,205			47.6
Pa		1,958	6,249	17,623			250.7
Xo		145,583	153,294	159,720			9.2
Co		43,675	46,087	47,916			9.2
Mco		10,919	11,776	11,979			9.0
Yo		101,908	107,207	111,804			9.2
So		74.520	85,744	94,130			22.9
Po		7,778	21,463	35,067			127.1
S		176,337	186,012	198,689			12.0
P'		61,408	70,648	79,962			26.2
P		23,937	33,177	42,491			55.9
Ch		123,436	130,539	140,243			12.9
If		131,318	136,425	143,166			8.7
Ifao		17,772	22,879	29,620			51.8
Mif		41,600	41,803	42,660			2.5
Y		251,362	256,661	261,258			3.9
Parameters							
lmif	1	1	1.005	1.025	2	99.5	2.5
lifao	1.5	1.5	1.931	2.5	2.5	51.8	51.8
lso	1.9	1.9	2.186	2.4	2.4	22.9	22.9
mif	0.25	0.291	0.306	0.325	0.6	114.3	11.1
mc	0.25	0.257	0.273	0.285	0.6	128.2	10.3
mco	0.25	0.25	0.256	0.274	0.6	136.7	9.4
sa	0.55	0.55	0.840	0.95	0.95	47.6	47.6
so	0.2	0.476	0.559	0.647	0.7	89.4	30.6
co	0.3	0.300	0.301	0.310	0.7	132.9	3.3
ch	0.7	0.700	0.702	0.715	0.95	35.6	2.1
ifao	0.8	0.800	0.826	0.964	—	—	19.9
tdh	0.03	0.032	0.035	0.037	0.06	85.7	13.2
tdao	0.03	0.42	0.057	0.080	0.25	386.0	13.3

between the minimum and maximum value for each of the variables. The latter types of averages cannot be used as central values, as they generally would not satisfy the identities and analytical constraints of the national accounts model.

Conclusions can also be drawn with regard to the direction in which national accounts estimates could be improved, once new sources of data become available or existing records are analysed in a more effective manner. Such improvements imply that the indirect measurement of selected national accounts items based on the lp algorithm is replaced by direct measurement of those items on the basis of newly available statistical information. Some of the effects of such replacement can be appreciated from the information included in Table 7. The first column of the table is the same as the last column of Table 6. It quantifies for each of

TABLE 7

Ex Post INTERVALS OF VARIABLES OF EXTENDED NATIONAL ACCOUNTS SCHEME AS COMPARED WITH REDUCED INTERVALS THAT RESULT ALTERNATIVELY FROM DIRECT MEASUREMENT OF IMPORTS (*Mco*, *Mif*) AND PRODUCTION ACCOUNT TRANSACTIONS OF OTHER INDUSTRIES (*Xo*, *Co*, *Yo*, *So*, *Po*)

	Unreduced *Ex Post* Intervals	Reduced intervals, based on direct measurement of: Imports	Production Account Transactions of Other Industries
Sa	47.6	52.0	22.9
Pa	250.7	173.2	91.1
Xo	9.2	7.3	—
Co	9.2	7.3	—
Mco	9.0	—	4.1
Yo	9.2	7.3	—
So	22.9	22.9	—
Po	127.1	130.0	—
S	12.0	11.4	3.9
P'	26.2	24.9	10.7
P	55.9	51.1	24.0
Ch	12.9	11.4	3.8
If	8.7	8.6	3.8
Ifao	51.8	48.5	25.4
Mif	2.5	—	1.1
Y	3.9	3.0	—

the variables the lengths of the *ex post* intervals as a percentage of their central values. The remaining two columns include, as illustrations, the reductions of those intervals as a result of introducing additional direct estimates of variables into the compilation procedure. Column 3 reflects direct estimation of all import categories (including *Mif* and *Mco*) on the basis of further detailed analysis of foreign trade data. Column 3 shows how the intervals are reduced, when direct estimates are made of all production account transactions of other industries with help of more comprehensive survey or census information covering all or a large part of the establishments that belong to this sector. As none of the improvements could be implemented in practice with the data available, it was assumed in the table that direct measurement resulted in estimates that were equal to the central values earlier calculated as presented in Table 6. The direct estimates in Table 7 are those that have zero intervals (—) in columns 2 and 3. The zero intervals also apply to those variables that could be derived as residuals, once the direct information became available. The latter holds for GDP (*Y*) in column 3 of the table, which can be derived once value added of other industries (*Yo*) is measured directly.

The information presented in columns 2 and 3 shows that direct estimates of imports or production account transactions of other industries reduce the *ex post* intervals of a majority of the variables, including those variables that are not estimated directly. What was to be expected is that the reduction in the intervals depends on the type of variables estimated directly. For instance, the table shows that there is much more impact from direct estimates of production account transactions of other industries (column 3) than of direct estimates of import categories (column 3). In the latter instance intervals are generally reduced a few percentage points only, while intervals are reduced to half or even less when direct estimates are made of production account transactions of other industries. The difference in impact between the two columns may give an indication of how to improve national accounts estimates in the future. If one has to choose between further studies of foreign trade records to improve the import estimates, or improved surveys of industries to expand the coverage of establishments in the other industries sector, the latter course clearly is preferable. The effects may be examined similarly for direct estimates of compensation of employees through social security records, direct estimates of private final consumption expenditure through household surveys, or improvement of gross fixed capital formation estimates through investment surveys of enterprises.

III. Evaluation and Further Development of the Modified Algorithm

This study is only a first step in the development of the lp algorithm for national accounts compilation. Much more work is needed to establish it more firmly from the conceptual point of view and make it operational beyond the application to the mere illustrative examples presented in section II. Some preliminary thoughts on some issues that need further elaboration, such as the theoretical justification for the central value concept, the practical and conceptual considerations in defining *ex ante* intervals for parameters and variables, and the required data processing facilities needed for the application of the algorithm in terms of the size of the linear programming matrix, are presented below. Finally some specific applications of the algorithm which might be considered in the field of national accounts as well as in other statistical systems are indicated.

The emphasis of the present study is on finding *ex post* intervals for variables and parameters which define the limits of the solution space. As no further information is available in the present model to arrive at optimal one point estimates, a central value vector was defined in the examples of section II as a simple unweighted average of all extreme solution vectors that were needed in order to arrive at the *ex post* minimum and maximum values. The advantage of this approach is that it is simple and easy to use. A more complex alternative is to apply the central value criteria defined in other recent similar studies [2, 3, 4] which are based on a modified version of the RAS method [5]. These approaches if used jointly with the present algorithm would minimize an objective function which might include the sum of the squared differences between the central values of all variables and parameters and their preliminary values before a reconciliation, weighted by the elements of the inverse of a covariance matrix. Such an approach, however, would complicate the use of the present algorithm because it would introduce non-linearities in the objective function and would require additional information on preliminary estimates in addition to the *ex ante* intervals that were used as the basis for the present algorithm. The incorporation of the covariance matrix furthermore seems to be superfluous in the present algorithm because the variances are already reflected in the restrictions of the present algorithm, while the covariances are difficult to measure and are generally assumed to be zero in the alternative approaches quoted.

The possibilities of establishing intervals for national accounts estimates [6, 7] have generally been limited for two reasons. The first is that a large part of the basic data used in national accounts estimation is a by-product of administrative records, for which it is difficult to assign reliability intervals on the basis of objective criteria. Only in instances where sample surveys are used as a source of statistical information can confidence intervals be determined with help of criteria developed in sampling theory. The second reason is a consequence of the present compilation procedures used in national accounts. Even if one could establish reliability intervals for the basic data, it would still be difficult to determine appropriate intervals for the ultimate national accounts estimates, given the very complex relation between these estimates and the basic data as a result of the adjustments applied to the latter. Through systematization of the national accounts compilation procedure based on the modified algorithm, the subjectivity of assigning reliability intervals can be reduced, however. The modified method accurately measures the ultimate *ex post* intervals of the variables in the national accounts framework, once the *ex ante* intervals of the basic data and the parameters of the analytical and adjustment relations have been defined. As this implies that the second obstacle mentioned above is no longer valid, one can concentrate on defining the reliability intervals of the basic data and particularly of the parameters of the analytical and adjustment constraints.

Three types of *ex ante* intervals could be used in the application of the algorithm:

(a) The first type are the intervals obtained from sample surveys. Probability functions and reliability intervals would thus be available for some of the national accounts variables and in some instances for parameters of the analytical relations.

(b) The second type of reliability intervals is based on judgemental criteria. Efforts have been made in some national accounting systems to assign subjective intervals to national accounts variables based on national accountants' experience [8, 9]. The proposed algorithm, however, opens the possibility not only to incorporate intervals for variables, but also for parameters in analytical functions, such as the ones used in the examples of section II. Such intervals may be established more easily, as national accountants continually make use of assumed ratios in the present compilation procedures and also make judgements on which ratio values are acceptable and which ones are not. Further development of the algorithm may require the identification of those parameters for which fairly short *ex ante* intervals can be set.

(c) Finally regression analysis may aid in defining intervals of parameters on the basis of a systematic analysis of relations between time series in the past. These intervals are generally indicated when analytical functions are measured. The proposed algorithm thus opens the possibility of incorporating the results of econometric analysis in the national accounts compilation procedure.

The size of the lp matrix is an important element in determining the data processing requirements and thus would influence the operationality of the lp algorithm. Experiments with the illustrative examples of section II have shown what the determining factors of the lp matrix size are. The number of columns of the matrix is equal to the number of unknown variables in the solution vector that are not equal to zero, plus one column for the solution vector itself. The number of variables that are not zero is equal to the total number of unknown variables minus the number of identities and analytical and adjustment constraints. To the number of unknown variables must be added the slack variables that are introduced in order to convert the analytical and adjustment inequalities to equalities required for the lp technique. In more precise terms, if the number of unknown variables (excluding slack variables) is X, the number of identitites is I, the number of analytical and adjustment constraints (pairs of inequalities) is S, then the number of columns of the lp matrix is equal to the number of unknowns including the slack variables $(X+2S)$ minus the number of identities and analytical constraints $(I+2S)$, plus 1 for the solution vector, or $(X-I+1)$. The number of rows of the matrix is equal to the number of national accounting identities and analytical constraints, or $(I+2S)$. Based on a number of 16 unknown variables represented in Table 3, and 4 identities and 13 adjustment and analytical relations shown in Table 4, the size of the lp matrix would thus be 13×30.

Further reductions of the size of the matrix, however, can be obtained in a number of ways. The number of unknown variables can be reduced by not including aggregates such as GDP and value added that can be defined as the sum of other unknown variables that can be included. This reduces not only the magnitude of X, but also the number of identities I because if an aggregate is included in the compilation structure, it also needs to be defined in terms of other component items that are included as unknown variables. Elimination of these aggregates therefore reduces the number of rows of the matrix, not the number of columns. Such reduction can be considerable, as is apparent from the example presented in section IIB, where the total number of variables including aggregates was 16 (in Table 3) and the number of unknown variables included in the national accounts model is only 10 (in Table 4). Another reduction in the size of the lp matrix can be obtained by removing from S all analytical and adjustment constraints that define one unknown national accounts category in terms of known basic data only. Such constraints, which can be reformulated as maxima and minima of the unknown variables, do not have to be incorporated in the matrix itself, but can be handled externally on the basis of the modified lp technique mentioned earlier (see Reference 1). This reduces S and therefore the number of rows of the matrix further. In the example of section IIB the number of 13 analytical and adjustment constraints is thus reduced to only 7 in the final formulation of the lp algorithm. In order to improve the operationality of the algorithm, it is essential that additional reductions of the lp matrix be achieved. Further study is therefore directed into the structure of the matrix in order to reveal dependencies between columns which would permit elimination of some of the rows and columns of the matrix, to identify parts of the matrix that could be processed in an iterative manner or independent parts of the matrix that could be processed separately.

The algorithm was applied in section II to compile production accounts for a limited number of sectors and to show a breakdown of GDP by cost components and expenditure categories. Its flexibility, however, allows it to be applied to any statistical system that is characterized by consistency requirements between the different types of statistical information. For instance, it could be applied to reconcile institutional sector data between different institutional sectors, or to input-output compilation supplementary to the earlier mentioned RAS method in order to eliminate some of the unacceptable results of that method. Furthermore, it could be used to reconcile information of statistical subsystems of financial flows and balance sheets, income distribution, balance-of-payments statistics and government finance statistics with related information of the national accounts. Reconciliation between micro data and macro national accounts aggregates could be another field of application. The uses are not necessarily restricted to national accounts only, but could be also directed to internally consistent statistical systems that cover demographic and social information.

REFERENCES

[1] Hadley, G. *Linear Programming*, Addison-Wesley, Reading, Mass. 1963, pp. 387-395.

[2] van der Ploeg, F. Reliability and the Adjustment of Sequence of Large Economic Acounting Matrices, *Journal of the Royal Statistical Society*, Series A (General), Volume 145, Part 2, 1982, pp. 169-194.

[3] Barker, Terry, Frederick van der Ploeg and Martin Weale, A Balanced System of National Accounts for the United Kingdom, *The Review of Income and Wealth*, Series 30, Number 4, December 1984, pp. 461-485.

[4] Arkhipoff, Oleg, The Synthesis and the Reliability of National Accounts by Means of Electronic Data Processing: Reflections on an Experience, published in *Models and Decision Making in National Economics*, Proceedings of the 2nd International Conference on Dynamic Modeling and Control of National Economies, North-Holland Publishing Co., 1979.

[5] United Nations, *Input-Output Tables and Analysis*, Series F, No. 14, Rev. 1, 1973, pp. 65-80.

[6] Morgenstern, Oskar, *On the Accuracy of Economic Observations*, Princeton, 1970, pp. 242-276.

[7] Arkhipoff, Oleg, Le Paradigme de la Mesure et la Fiabilité de la Compatibilité Nationale, *Journal de la Societé de Statistique de Paris*, ler trimestre, 1984, pp. 25-41.

[8] Central Statistical Office, *National Accounts Statistics*, London 1968.

[9] Australian Bureau of Statistics, *The Accuracy and Reliability of the Quarterly Australian National Accounts*, Occasional Paper No. 1982/2.

Erratum

Professor Marcel Das has kindly brought to my attention that there is a misprint on page 31 of Chapter 4, in the fourth definitional relationship in Table 1.

Where is says:

$$\bar{C}g = \text{ Governmental final consumption expenditure (37)}$$

it should read:

$$\bar{C}g = \text{Governmental final consumption expenditure (30).}$$

Chapter 5

National accounts estimation using indicator ratios

Review of Income and Wealth
Series 46, Number 3, September 2000

NATIONAL ACCOUNTS ESTIMATION USING INDICATOR RATIOS

BY JAN R. MAGNUS

CentER, Tilburg University

JAN W. VAN TONGEREN

United Nations Statistics Division

AND

AART F. DE VOS

Free University, Amsterdam

We propose a new approach to national accounts compilation, which also serves as a formalization of current compilation practices. When formalizing the procedure, a distinction is made between (basic) data, national accounts identities and so-called indicator ratios. The latter are ratios of or percentage relations between national accounts variables, such as the relation between output and value added. Indicator ratios are currently used in national accounts compilation practices in order to make adjustments to the basic data or to fill in missing data. The latter use is particularly relevant when basic data are scarce, which is the case not only in many developing countries, but also in developed countries when annual accounts are compiled for recent periods. The (basic) data, indicator ratios and identities together are used in a Bayesian approach to estimate the values of national accounts variables and analytical indicator ratios based thereon. The amendment of the current practices consists in introducing reliability intervals of basic data and indicator ratios, which allows for the use of a much larger number of indicator ratios in the compilation and checking of national accounts data. The Bayesian compilation approach makes it possible—in contrast to current practices—to use indicator ratios both as priors and as analytical indicators.

1. INTRODUCTION

This paper proposes a new and relatively simple Bayesian approach to national accounts compilation. The new approach can (and will) be compared with current compilation practices, which are generally less sophisticated and do not provide reliability intervals of the estimates. The proposed approach takes account of the identities and the indicator ratios used in the national accounts compilation, and establishes a direct link to an important usage of national accounts figures in analysis, namely the use of indicator ratios. It also provides a framework for evaluating the usefulness of extra or better data, for example in estimating and analyzing indicator ratios.

Current national accounts compilation practices use basic data, identities (reflecting internal consistency criteria that the national accounts estimates should satisfy), and structural coefficients (or indicator ratios). The latter are ratios of

Note: We are grateful to Sanne de Boer and Bernhard Conlon for help with the original programming in Gauss, to Anton Markink for professional help in preparing robust computer packages, and to the participants of the "Expert Group Meeting on the Use of Macro Accounts in Policy Analysis," 5–9 October 1998, at the United Nations in New York, and two referees for helpful and constructive comments.

or percentage relations between national accounts variables, such as the relation between output and value added, between value added of large and small establishments, or between product flows and trade-and-transport margins. The indicator ratios are currently used in order to make conceptual adjustments to the basic data, or adjustments for coverage of basic data, to derive missing data, and to perform consistency checks on the data. Indicator ratios are particularly relevant in the compilation of national accounts when basic data are scarce. This is the case in many developing countries, but also in developed countries when annual accounts are compiled for recent periods or quarterly accounts are compiled.

The current paper provides a formalization of current practices, in line with recent UNSD attempts (United Nations, 1999) which support the increasing computerization of national accounts compilation procedures in current practices and link those practices to methodologies of measurement used in analyses based on the national accounts estimates, in particular those used in econometrics.

The discussion in the paper focuses on the national accounts as the medium through which the interaction between indicator ratios used in compilation and analysis is studied. National accounting in this context should be interpreted in a very broad sense, i.e. an approach to integrate and make consistent basic data through the use of an accounting framework with a consistent set of concepts and classifications. This is not the limited concept of national accounts reflected in a large number of country practices, which focuses on the estimation of GDP only. The accounting framework referred to here is that of the international guidelines of the 1993 System of National Accounts (hereafter SNA) described in United Nations *et al.* (1993), which emphasize—in addition to GDP—the compilation of so-called institutional sector accounts and also introduces the possibility of using non-economic data in the compilation of satellite accounts. However, since the example in Section 5 uses only economic data, national accounts in the present context refers to national economic accounts only.

Indicators are widely used by international organizations (World Bank, 1993; United Nations Development Program, 1996) to set policy goals and monitor development in countries and regions. Indicators in this broad sense include basic data, estimates that are reconciled through national accounting, and also ratios between variables. The present paper deals only with indicator ratios. The reason for this limitation is twofold. First, indicator ratios, such as per capita GDP or investments as a percentage of GDP are more relevant for international and intertemporal comparisons than the underlying variables of GDP, investments and population size. Indicator ratios can thus be considered as an important analytical summary of a large national accounting data set. Secondly, indicator ratios are also used in the compilation of national accounts and it is the interaction between these two uses which is the main focus of this paper.

The paper also contributes to an old and still very relevant discussion of reliability of basic data, for which Morgenstern (1963, first published in 1950) laid the foundations. Morgenstern approached national accounting as a branch of descriptive statistics. This is different from the approach taken in this paper, where national accounting is used as the medium for integration of statistics. The approach initiated by Morgenstern was followed by at least three countries which publish their national accounts estimates and related statistics with indications of their reliability: the U.K., Australia and more recently Canada. The reliability of national accounts can also be based on revisions made over time, an approach already suggested by Morgenstern.

A second approach was initiated by Stone, Champernowne and Meade (1942). This involves both the reliability of the data and the problem of balancing the accounts, see Byron (1978); Barker, Van der Ploeg and Weale (1984); and Van der Ploeg (1982, 1984, 1985).

Another approach was followed by van Tongeren (1985), who used linear programming techniques not only to balance the accounts, but also to determine the relation between the "prior" reliability intervals of basic data and the "posterior" intervals, once those data are reconciled within an accounting framework.

The present paper builds on these approaches. It is written both for econometricians and other theoretical statisticians, familiar with estimation procedures used in econometrics and related fields, as well as for national accountants, familiar with the estimation methods used in national accounting. The integration of the two approaches might be beneficial to both disciplines, provided national accountants pay more attention to the analyses in which the data are used, and econometricians familiarize themselves more with the intricacies of basic data. National accounting, which was initially designed for analytical purposes, but was later developed by statisticians, may be the ideal framework for such an interdisciplinary approach.

The plan of the paper is as follows. In Section 2 we state our main theoretical result (proved in an Appendix), which gives us the means of combining incomplete data with incomplete prior information (including exact accounting identities). This theorem is based on normality and linearity. In Section 3 we discuss how to use indicator ratios—which are nonlinear—as priors. This leads to an extension of Theorem 1. In Section 4 we discuss how to deal with multiple priors, that is the situation where many priors or several priors on the same variable are available. Section 5 contains a simple but realistic example of the application of the proposed approach, and Section 6 concludes. The proof of our main theorem is provided in the Appendix.

2. The Adjustment of Unreliable Observations

Consider a vector x of n latent variables, to be regarded as a vectorization of a system of accounts. Data are available on $p \leq n$ components (or linear combinations) of x. Let d denote the $p \times 1$ data vector. Our starting point is a *measurement equation*, $d \mid x \sim N_p(Dx, \Sigma)$. Typically, the $p \times n$ matrix D is a selection matrix, say $D = (I_p, 0)$, so that Dx is a subvector of x. Measurements are unbiased in the sense that $E(d \mid x) = Dx$. The $p \times p$ matrix Σ denotes a positive definite variance matrix, typically (but not necessarily) diagonal.

In addition to the p data, we have access to two further pieces of information: deterministic accounting constraints and prior (possibly multiple) views concerning the latent variables or linear combinations thereof. These two pieces are combined into one set of linear *priors*, $Ax \sim N_m(h, H)$, where the variance matrix H is singular, because each of the deterministic constraints has variance 0.

We now have data and priors, and we wish to employ Bayes' theorem to combine them and obtain posteriors. We emphasize two complicating features of this problem: we do not, in general, have data on all latent variables ($p<n$), nor do we have priors on all latent variables ($m<n$). If we had data on all latent variables ($p=n$), then the results in Van der Ploeg (1985, Section 2.1) could be applied. If we had priors on all latent variables ($m=n$), then Lemma A1 in the Appendix would give the desired posterior distribution. It is the joint occurrence of lacking data and insufficient prior information, which makes the problem difficult.

Clearly we need an identifiability condition, since each latent variable must be revealed either through the data or through the priors or both. A necessary and sufficient identifiability condition, easy to check, is given in Theorem 1, which also provides a complete solution to the general problem discussed above.

Theorem 1. Let x be an $n\times 1$ vector of latent variables and let d be a $p\times 1$ data vector such that

$$d|x\sim N_p(Dx,\Sigma), \tag{1}$$

where the $p\times n$ matrix D has full row-rank and Σ is positive definite (hence nonsingular). Suppose that prior information is available in the form

$$Ax\sim N_m(h,H), \tag{2}$$

where the $m\times n$ matrix A has full row-rank and H may be singular. If $m<n$, let L be a semi-orthogonal $n\times(n-m)$ matrix (that is, $L'L=I_{n-m}$) such that $AL=0$, and assume that the identifiability condition

$$r(A)+r(DL)=n \tag{3}$$

is satisfied. Then the posterior distribution of x is given by

$$x|d\sim N_n(\mu,V) \tag{4}$$

with

$$V=A^+HA^{+\prime}-A^+HA^{+\prime}D'\Sigma_0^{-1}DA^+HA^{+\prime}+CKC' \tag{5}$$

and

$$\mu=A^+h-(A^+HA^{+\prime}+CK)D'\Sigma_0^{-1}(DA^+h-d), \tag{6}$$

where

$$\Sigma_0=\Sigma+DA^+HA^{+\prime}D',\qquad A^+=A'(AA')^{-1},\qquad C=I_n-A^+HA^{+\prime}D'\Sigma_0^{-1}D, \tag{7}$$

and

$$K=\begin{cases}L(L'D'\Sigma_0^{-1}DL)^{-1}L', & \text{if } m<n,\\ 0, & \text{if } m=n.\end{cases} \tag{8}$$

Proof. See Appendix. ||

Let us provide a very simple example in order to demonstrate application of the theorem. In this example we have one national accounts identity $y=c+i+g$

and two independent data: $y=230$ and $g=-44$. In addition, our prior belief is that c should be around 220.5 and i/c around 1/3. We do not yet know how to treat ratios (this is discussed in the next section), but we can deal with this situation naively as follows:

naive example:

data:	$y=230$ (11.5)	*priors*:	$y=c+i+g$
	$g=-44$ (2.2)		$c=220.5$ (11.025)
			$i=73.5$ (3.675).

Standard errors are given in parentheses. In each case the coefficient of variation (standard error divided by mean) is assumed to be 5 percent. We have $n=4$, $p=2$, and $m=3$. The vector of latent variables is $x=(y,c,i,g)'$, and

$$D=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix},\qquad \Sigma=\begin{pmatrix}(11.5)^2 & 0\\ 0 & (2.2)^2\end{pmatrix},\qquad h=\begin{pmatrix}0\\ 220.5\\ 73.5\end{pmatrix},\tag{9}$$

$$A=\begin{pmatrix}1 & -1 & -1 & -1\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix},\qquad H=\begin{pmatrix}0 & 0 & 0\\ 0 & (11.025)^2 & 0\\ 0 & 0 & (3.675)^2\end{pmatrix}.\tag{10}$$

The semi-orthogonal matrix L is $(1/\sqrt{2})(1,0,0,1)'$ and the identifiability condition is satisfied. Application of Theorem 1 then gives the following results:

posterior moments:	$y=239.7$ (8.2)
	$c=211.6$ (8.2)
	$i=72.5$ (3.6)
	$g=-44.4$ (2.2).

Clearly the accounting identity is exactly satisfied. One can experiment with this simple example to see how different assumptions on the precisions of data and priors affect the results. Equally, we could allow for correlations between priors or measurements (data).

It is easy to see that the identifiability condition in Theorem 1 is the most general possible—it is necessary and sufficient.

In classical statistics, it is well-known that there exists a close and non-trivial connection between generalized least squares and best linear unbiased estimation; see Rao (1973, pp. 294–302) and Magnus and Neudecker (1999, Section 13.18). An analogous result holds in Bayesian statistics. If we partition the prior information $Ax\sim N_m(h,H)$ into two parts:

$$\pi_1: A_1x\sim N_{m_1}(h_1,H_1),\qquad \pi_2: A_2x=h_2\quad \text{(a.s.)},\tag{11}$$

where h_1 has m_1 components, h_2 has $(m-m_1)$ components, and H_1 is positive definite, then a generalized least squares procedure would minimize

$$(d-Dx)'\Sigma^{-1}(d-Dx)+(A_1x-h_1)'H_1^{-1}(A_1x-h_1)\tag{12}$$

subject to the linear constraints

$$A_2 x = h_2. \tag{13}$$

Since (12) is proportional to the exponential term of the posterior when only π_1 is used, while also using the linear constraints (13) is equivalent to adding π_2, we see that minimization leads to the posterior mode, which, in view of the normality assumption, equals the posterior mean $x=\mu$ of Theorem 1. Hence, there exists a close connection between our Bayes solution and generalized least squares; see also Van der Ploeg (1985).

Analogous to the classical result that a generalized least squares estimator is best linear unbiased in the absence of the normality assumption (by the Gauss–Markov theorem), there is a Bayesian result that μ is a Linear Bayes estimator, obtained by minimizing the expected (with respect to data) quadratic loss in the class of linear functions of the data. This result does not depend on normality assumptions, only on first and second moments. As our posterior mean is linear in d, it may be justified as a Linear Bayes estimator; see e.g. Goldstein (1988). O'Hagan (1994, pp. 163–66) gives a short and critical review of Linear Bayes estimators.

Note that other assumptions than normality, like more heavy-tailed priors, lead to Bayesian posterior means that are not linear in the data. Numerical methods like the Gibbs sampler are then required to estimate the posteriors. However, the Linear Bayes estimator may still be used as a near-optimal simple device in such cases.

3. Indicator Ratios

In practice many of the priors will be nonlinear. In particular, many of the priors used in the construction of national accounts are "indicator ratios," that is, ratios of two latent variables. In this section we shall see how indicator ratios can be linearized in a suitable manner, so that Theorem 1 can still be applied.

Let x and y be latent variables and consider an indicator ratio $R=y/x$. Let r denote the prior expectation of R. Assume for the moment that we have prior information on R and x and that these priors are independent. This will not be strictly true in practice, but is nevertheless reasonable in many applications. Then we can easily show that

$$E(y-rx)=0, \qquad \mathrm{var}(y-rx)=\mathrm{var}(R)\cdot(\mathrm{var}(x)+(Ex)^2). \tag{14}$$

Our strategy is to replace the prior R by its linearization $y-rx$. We have prior knowledge about the mean and variance of R, but *not* about the mean and variance of x. If we would know the mean and variance of x, then we could replace the prior on R by a prior on $y-rx$, and Theorem 1 could be applied. Since we do not know the moments of x, we use a simple iterative procedure, as follows. First, use (14) taking Ex to be the value of x in the previous year and let $\mathrm{var}(x)=0$. Then, apply Theorem 1. This gives posterior moments of all latent variables, and hence in particular of x. In step 2 we use the posterior moments Ex and $\mathrm{var}(x)$ obtained in step 1 and recalculate the prior variance of $y-rx$ from (14).

Using this updated prior variance, we apply Theorem 1 again and continue this process until convergence.

We may or may not have access to last year's values. This is not important, since the resulting posterior estimates will be independent of the starting values of the iteration. In practice we will have not one but several indicator ratios. The iteration procedure is then applied to all of them simultaneously.

The linearization just described involves an approximation. Our experience shows that the approximation works well in practice, but of course there may be situations where it does not work well. Below we provide a further justification for the procedure.

Consider a very simple set-up with two latent variables x and y and one indicator ratio $R = y/x$, as follows:

$$\text{(15)} \qquad \textit{data}: \quad d \mid x \sim (x, \sigma^2)$$

$$\text{(16)} \qquad \textit{prior } \pi: \quad R \sim (r, \tau^2).$$

We wish to replace the prior π in (16) by its linearization π',

$$\text{(17)} \qquad \textit{prior } \pi': \quad y - rx \sim (0, \tau_*^2),$$

and we wish to choose τ_*^2 optimally in some sense. Given (15) and (17), the posterior moments of y given d can be calculated from Theorem 1 or directly. They are

$$\text{(18)} \qquad y \mid d \sim (rd, r^2\sigma^2 + r_*^2).$$

Given (15) and (16), what are the posterior moments of $y \mid d$? The prior π directly implies that

$$\text{(19)} \qquad E(y \mid x) = rx, \qquad \operatorname{var}(y \mid x) = \tau^2 x^2.$$

Let $z = x \mid d$. Then the posterior moments of $y \mid d$ are

$$\text{(20)} \qquad \begin{aligned} E(y \mid d) &= E_z E(y \mid d, z) = E_z E(y \mid d, x) \\ &= E_z E(y \mid x) = E_z(rx) = rd \end{aligned}$$

and

$$\text{(21)} \qquad \begin{aligned} \operatorname{var}(y \mid d) &= E_z \operatorname{var}(y \mid d, z) + \operatorname{var}_z E(y \mid d, z) \\ &= E_z \operatorname{var}(y \mid x) + \operatorname{var}_z E(y \mid x) \\ &= E_z(\tau^2 x^2) + \operatorname{var}_z(rx) \\ &= \tau^2(d^2 + \sigma^2) + r^2\sigma^2. \end{aligned}$$

We see that the posterior mean of $y \mid d$ is rd, both in the linearized and the non-linearized version. We now choose the prior variance τ_*^2 such that also the posterior variance of $y \mid d$ for the linearized prior (17) equals that for the non-linearized prior (16). This yields

$$\text{(22)} \qquad \tau_*^2 = \tau^2(d^2 + \sigma^2).$$

Hence, in this simple example, the iterative procedure is justified by the fact that it leads to the correct first two posterior moments. We notice that this method is not Linear Bayes, since the estimator is not linear in d. A Linear Bayes estimator would have a larger variance.

Continuing our simple example from the previous section, let us assume, instead of $i=73.5$, that $i/c=1/3$ or alternatively that $c/i=3$. The results are presented in Table 1.

TABLE 1
POSTERIOR MEANS AND STANDARD ERRORS FOR THREE PRIOR SPECIFICATIONS OF THE SIMPLE MODEL

	Prior		
Posterior	$i=73.5$	$i/c=1/3$	$c/i=3$
y	239.7 (8.2)	237.2 (9.2)	237.2 (9.2)
c	211.6 (8.2)	211.6 (7.1)	211.6 (7.0)
i	72.5 (3.6)	69.9 (3.7)	69.9 (3.7)
g	−44.4 (2.2)	−44.3 (2.2)	−44.3 (2.2)

The difference between the three specifications is very small. The two alternative prior indicator ratios ($i/c=1/3$ and $c/i=3$) do not necessarily yield the same posterior estimates, because different approximations are involved in linearizing the ratios. Nevertheless, the difference between $i/c=1/3$ and $c/i=3$ as prior indicator ratio is negligible.

We conclude that an iterative version of Theorem 1 can be suitably applied to linearized indicator ratios in order to obtain posterior moments of all latent variables in the system.

4. MULTIPLE PRIORS

There is one further possible complication. We may have many priors or several priors on the same latent variable. In such situations the $m \times n$ matrix A in the prior specification

$$Ax \sim N_m(h, H) \tag{23}$$

will have rank $r<m$ and the conditions of Theorem 1 are not satisfied. For example, if $y=c+i$, we may have priors on both c/y and i/y, and these will generally not add up to one. In such situations we will need to reflect carefully about the source of the conflict. Also it may occur that such priors are *not* in conflict. Then we need to assess how dependent the information is. Maybe the information on c/y and i/y originated from the same source, in which case they are perfectly correlated. Much has been written about combining expert's opinions; see Genest and Zidek (1986); Wiper and French (1995); and Clemen and Winkler (1999).

The most common approach to this problem uses different expert opinions like "data." We follow this approach and generalize it to a format that suits our goals. Let S be an $m\times(m-r)$ matrix and T an $m\times r$ matrix such that (S, T) is orthogonal,

$$(24) \qquad S'S=I_{m-r}, \qquad T'T=I_r, \qquad S'T=0,$$

and satisfies $A'S=0$. (Hence S contains the eigenvectors associated with the $m-r$ zero eigenvalues of AA' and T contains the remaining r eigenvectors.) Premultiplying (23) by $(S, T)'$ gives the equivalent prior specification

$$(25) \qquad \begin{pmatrix} S'Ax \\ T'Ax \end{pmatrix} \sim N_m\left[\begin{pmatrix} S'h \\ T'h \end{pmatrix}, \begin{pmatrix} S'HS & S'HT \\ T'HS & T'HT \end{pmatrix}\right],$$

and hence $T'Ax|(S'Ax=0)\sim N_r(h^*, H^*)$ with

$$(26) \quad h^*=T'h-T'HS(S'HS)^+S'h, \qquad H^*=T'HT-T'HS(S'HS)^+S'HT,$$

where $(S'HS)^+$ denotes the Moore–Penrose inverse of $S'HS$. Instead of the prior information on Ax contained in (23) we now consider the prior information on $T'Ax$. The matrix $T'A$ has full row-rank and thus satisfies the conditions of Theorem 1.

A trivial example may illustrate this procedure. Suppose we have two pieces of information on a single latent variable x,

$$(27) \qquad x\begin{pmatrix} 1 \\ 1 \end{pmatrix} \sim N\left[\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \tau^2 \end{pmatrix}\right].$$

According to one expert, $Ex=1$; according to another expert with independent information, $Ex=2$. Then,

$$(28) \qquad S=(1/2)\sqrt{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad T=(1/2)\sqrt{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Substituting into (23) gives

$$(29) \qquad x\sim N(2-\lambda^2, \lambda^2), \qquad \lambda^2=\tau^2/(1+\tau^2),$$

the standard weighted average of the two pieces of information.

5. An Example

We illustrate the proposed methodology by investigating a simple but realistic accounting framework with data based on a Western-type economy. The framework includes national aggregates, such as GDP, expenditures and external transactions, that allow for the derivation of national disposable income and national saving (not explicitly included in the presented accounts). It also includes the main institutional sector accounts, as published in the 1993 System of National Accounts (United Nations *et al.*, 1993).

The framework and data, involving 38 variables distributed over 5 sectors, is presented in Table 2 together with last year's data. The framework and the data are described in detail in United Nations (2000, Chapter 4A) and are consistent with the international guidelines of the 1993 System of National Accounts (SNA) described in United Nations *et al.* (1993).

TABLE 2

NATIONAL ECONOMIC ACCOUNTS, EXAMPLE WITH LAST YEAR'S DATA

Total Economy		Resident Sectors		
Industries	Rest of the World	Government	Non-financial and Financial Corporations	Households
[1] Output, incl. product taxes, less subsidies 3,737	[10] Imports 499			
[2] Intermediate consumption 1,883	[11] Exports 540	[17] Final consumption, government 368		[26] Final consumption, housholds 1,031
[3] Gross capital formation, total economy[1] 414		[18] Gross capital formation, government[1] 40	[22] Gross capital formation, corporations[1] 287	[27] Gross capital formation, households[1] 87
[4] GDP, market prices (current prices) 1,854 [5] GDP, market prices (constant prices) 1,160	[12] External balance of goods and services 41			[28] Disposable income, gross[5] 1,259 [29] Disposable income before taxes, gross[6] 1,437 [30] Savings, gross 228
[6] Compensation of employees, paid and mixed income, gross 1,204	[13] Compensation of employees received by residents, less paid to non-residents 4			[31] Compensation of employees and mixed income received 1,208
[7] employment (1,000 m/years worked) 33,350				[32] Size of population (× 1000), beginning of year 88,700 [33] Size of population (× 1000), end of year 90,000 [34] Population increase (× 1000) 1,300
[8] Taxes on production and imports, less subsidies[2] 191	[14] Taxes on production less subsidies plus taxes on income and wealth, received by resident government less paid to non-resident government 1	[19] Taxes on production less subsidies plus taxes on income and wealth, received by government 404	[23] Taxes on income and wealth, paid by corporations 34	[35] Taxes on income and wealth, paid by households 178

TABLE 2—continued

Total Economy		Resident Sectors		
Industries	Rest of the World	Government	Non-financial and Financial Corporations	Households
[9] Operating surplus, gross (excl. mixed income) 459	[15] Other incomes, receipts by residents less payments to non-residents[3] −8	[20] Other outlays, payments less receipts, by government[4] 46	[24] Other incomes, receipts less payments, by corporations[3] 257	[36] Other incomes, receipts less payments, by households[3,7] 229 [37] Capital transfers, receipts less payments, households 11
	[16] Net lending to abroad 38	[21] Net lending, government −50	[25] Net lending, corporations −64	[38] Net lending, households 152

Note: All entries, except [7] and [32]–[34], are in millions of US$.

[1]Gross capital formation includes the value of improvements to land and the cost of ownership transfers of non-produced assets.

[2]Production taxes less subsidies have not been allocated to sectors, but have only been recorded for the total economy.

[3]Other incomes, receipts less payments, include operating surplus gross, property income, and non-tax current and capital transfers. Capital transfers include acquisition less disposal of non-produced non-financial assets.

[4]In the case of the government, other incomes have been replaced by other outlays, which include payments less receipts of property income and non-tax current and capital transfers, less operating surplus gross.

[5]Includes adjustment for the change in net equity of households on pension funds, and is after deduction of taxes on income and wealth.

[6]No tax deductions have been made.

[7]In the case of households, operating surplus excludes mixed income and capital transfers, which are presented separately.

The columns of the Table 2 refer to sectors of the economy and the rows to accounts. The first column contains the aggregate data on industries, while the other columns refer to the rest of the world and three aggregate resident institutional sectors, i.e. government, non-financial and financial corporations, and households. The rows of accounting data for each sector are grouped together by four accounting segments. The first segment, contained in the first three row-blocks, refers to the data elements of the supply and use table, i.e. output, imports, exports, intermediate and final consumption, and capital formation. The second segment, covering row-blocks 4 and 6, refers to the main product, income and related aggregates including GDP at current and constant prices, disposable income before and after taxes, saving, and the corresponding employment and population data needed to derive product and income aggregates per worker and per capita. The third segment, covering row-blocks 5, 7 and 8, refers to receipts and payments of compensation of employees including mixed income, taxes, operating surplus and other income and outlay data. The fourth segment, contained in row-block 9, includes net lending for each sector. This is used as the main analytical balancing item in each sector, except in the household sector where also disposable income and savings are shown.

The 38 variables of the accounting framework must satisfy 16 linear accounting restrictions: 10 "vertical" restrictions that define the accounting constraints within each sector, and 6 "horizontal" restrictions that correspond to accounting restrictions between the sectors. These are given in Table 3. In fact, there is one further "horizontal" restriction referring to net lending between sectors: [16] = [21] + [25] + [38], but this identity is linearly dependent on the other restrictions.

TABLE 3

ACCOUNTING IDENTITIES

Industries	[4] = [1] − [2]
	[9] = [4] − [6] − [8]
Rest of the world	[12] = [11] − [10]
	[16] = [12] + [13] + [14] + [15]
Government	[21] = [19] − [17] − [18] − [20]
Corporations	[25] = [24] − [23] − [22]
Households	[30] = [28] − [26]
	[29] = [28] + [35]
	[36] = [28] − [31] + [35]
	[38] = [30] + [37] − [27]
Supply and use	[1] + [10] = [2] + [3] + [11] + [17] + [26]
Gross capital formation	[3] = [18] + [22] + [27]
Compensation of employees	[6] + [13] = [31]
Population	[34] = [33] − [32]
Taxes	[19] = [8] + [14] + [23] + [35]
Other incomes and outlays	[24] + [36] + [37] − [9] − [15] − [20] = 0

An essential ingredient in estimating the 38 latent variables is prior, but uncertain, knowledge about some indicator ratios. We have selected, on economic grounds, 16 indicator ratios on which we believe reasonable prior knowledge to be available. These were selected from a larger list of indicator ratios presented in United Nations (2000, Chapter 4A). The 16 indicator ratios, together with their prior moments, are presented in Table 4.

Each indicator ratio except one has a prior mean which is assumed to be equal to last year's value, that is, our prior belief is that the indicator ratio remains the same. The only exception is the inflation rate, which we give a prior value of 2 percent, based on the results of a limited consumer price survey. As a consequence, the GDP price deflator [4]/[5] has a prior mean of last year's value multiplied by 1.02. We emphasize that the choice of prior means based on last year's values is *not* demanded by some theory. In the presence of more insights into the economy, other prior means could be used.

The assumed precision (uncertainty) of the indicator ratios is given as high (H), medium-high (MH), medium-low (ML), or low (L). These four categories indicate the coefficient of variation, that is, standard error divided by mean. In particular, "H" indicates a coefficient of variation of 0.5 percent, "MH" of 2.5 percent, "ML" of 5 percent, and "L" of 10 percent. For example, the first indicator ratio [3]/[4] has a prior mean of 0.2233 and a coefficient of variation of 2.5 percent. Hence the standard error is $0.025 \times 0.2233 = 0.0056$. In our example, the precision of the indicator ratios is "H," "MH" or "ML," but never "L."

In addition to the indicator ratios, we have four "other" priors, namely variables which we assume a priori will not change compared to last year. These are given in Table 5.

TABLE 4

INDICATOR RATIOS WITH PRIOR MOMENTS

			Indicator Ratios	Prior Moments: Mean	Prior Moments: Coefficient of Variation	Prior Moments: Standard Error
Indicator ratios vertically defined within industries and sectors	Production by industries	[3]/[4]	Investment share in GDP	0.2233	MH	0.0056
		[4]/[1]	Value added/output coefficient, total	0.4961	H	0.0025
		[4]/[5]	GDP price deflator	1.6302	H	0.0081
		[5]/[7]	GDP constant prices per worker, labor productivity (1,000 US$ per m/year)	0.0348	MH	0.0009
		[6]/[4]	Labor share in GDP	0.6494	H	0.0032
		[8]/[4]	Production taxes less subsidies/GDP	0.1030	MH	0.0026
	Corporations	[23]/[24]	Taxes/revenues of corporations	0.1323	MH	0.0033
		([24] − [23])/[22]	Earnings (after taxes)/gross capital formation, corporations	0.7770	MH	0.0194
	Households	[26]/[28]	Propensity to consume of households	0.8189	MH	0.0205
		[27]/[30]	Capital formation/saving, households	0.3816	MH	0.0095
		[31]/[29]	Labor income as share of disposable income of households, before taxes	0.8406	MH	0.0210
		[35]/[29]	Tax ratio of household disposable income, before taxes	0.1239	MH	0.0031
		[34]/[32]	Population growth	0.0147	ML	0.0007
Indicator ratios horizontally defined across industries and sectors		[10]/([1] + [10])	Import/supply-use	0.1178	MH	0.0029
		[17]/[26]	Government/household consumption ratio	0.3569	MH	0.0089
		[7]/[33]	Number of employees/population	0.3706	MH	0.0093

TABLE 5

OTHER PRIORS USED

Variable	Prior Moments		
	Mean	Coefficient of Variation	Standard Error
[13] Compensation of employees received by residents, less paid to non-residents	4	L	0.4
[14] Taxes on production less subsidies plus taxes on income and wealth, received by resident government less paid to non-resident government	1	L	0.1
[15] Other incomes, receipts by residents less payments to non-residents	−8	L	0.8
[37] Capital transfers, receipts less payments, households	11	L	1.1

Tables 3–5 give the accounting restrictions, the indicator ratios, and the "other" priors, 36 priors all together. However, we also have data. In fact we have data on 19 of our 38 variables. These are given, together with their assumed precisions (measurement errors), in Table 6.

Of course, the coefficient of variation is not defined when the mean is zero, and is not very useful when the mean is "close" to zero. Here we let the standard error be 0.5 when the mean (in absolute value) is smaller than 10.

Thus, we have 38 variables ($n = 38$), 19 data ($p = 19$), and 36 priors ($m = 36$) of which 16 identities. Using the iterative version of Theorem 1 on linearized indicator ratios (see Section 3), we obtain posterior means and variances of all 38 latent variables. Of the total number of 38 variables, we selected 11 "key" variables and 11 "key" indicator ratios. These are considered to be the essential elements for assessing the state and development of the economy under investigation.

5.1. *Key Variables*

Bayesian estimates (together with their standard errors) of key variables are presented in Table 7 (under "comprehensive compilation"). The key variables include, among others, GDP at market prices, gross capital formation, disposable income of households, and net lending of all sectors, including corporations, government, households and abroad. Some of the key variables can be estimated very accurately: gross capital formation in the total economy [3], GDP [4], final consumption government [17], final consumption households [26], and gross disposable income [28]. However, some of the key variables, in particular the balancing items, are estimated less accurately, as was to be expected. These include net lending by government [21], corporations [25] and households [38], and also gross household savings [30]. The most difficult to estimate are two key variables relating to the rest of the world: external balance of goods and services [12] and net lending to abroad [16]. Qualitatively these results appear to be sensible. The power of our approach, however, lies in the quantification of estimates and precisions.

TABLE 6

THE DATA AND THEIR PRECISIONS

Data	Moments		
	Mean	Coefficient of Variation	Standard Error
[1] Output, incl. product taxes, less subsidies	4,034	ML	201.7
[3] Gross capital formation, total economy	449	L	44.9
[6] Compensation of employees, paid and mixed income, gross	1,252	MH	31.3
[7] Employment (1000 m/years worked)	33,657	MH	841.4
[10] Imports	543	MH	13.6
[11] Exports	567	MH	14.2
[13] Compensation of employees received by residents, less paid to non-residents	3	*	0.5
[14] Taxes on production less subsidies plus taxes on income and wealth, received by resident government less paid to non-resident government	0	*	0.5
[15] Other incomes, receipts by residents less payments to non-residents	−10	ML	0.5
[17] Final consumption, government	385	MH	9.6
[18] Gross capital formation, government	41	MH	1.03
[19] Taxes on production less subsidies plus taxes on income and wealth, received by government	366	MH	9.2
[20] Other outlays, payments less receipts, by government	48	MH	1.2
[22] Gross capital formation, corporations	316	ML	15.8
[23] Taxes on income and wealth, paid by corporations	29	ML	1.45
[24] Other incomes, receipts less payments, by corporations	275	ML	13.75
[26] Final consumption, households	1,070	L	107.0
[32] Size of population (× 1,000), beginning of year	90,000	MH	2,250
[33] Size of population (× 1,000), end of year	91,200	MH	2,280

Note: When the prior mean (in absolute value) is smaller than 10, we set the standard error at 0.5.

We now compare our Bayesian estimates with the estimates based on a formalization of current practices (hereafter "SNA estimates"). These current practices do not involve the use of reliability intervals, and as a consequence can accommodate fewer structural coefficients (indicator ratios). The SNA estimates were obtained by simulating current compilation practices. The compilation is carried out in two stages. The first stage is based on the 19 data (the same 19 data that are used in the Bayesian approach). To "estimate" the 38 variables, we add the 16 restrictions (the same restrictions that we used in the Bayesian approach) and a small number (5 in this case) of the available indicator ratios. The five selected indicator ratios are those that are closest to the ones used in current practices, and consist of value added/output coefficient, total; the GDP price deflator; GDP in constant prices per worker, labor productivity; production taxes less subsidies/GDP; and also capital formation/saving of households (see

TABLE 7

BAYESIAN ESTIMATES FOR "KEY" VARIABLES

Key Variable	Previous Year	Comprehensive Compilation: Current SNA Practice	Comprehensive Compilation: Bayes	Reduced Compilation: Bayes
[3] Gross capital formation, total economy	414	490	424 (8.8)	434 (11.2)
[4] GDP, market prices (current prices)	1,854	2,001	1,890 (22.9)	1,938 (30.8)
[12] External balance of goods and services	41	25	29 (13.1)	43 (15.0)
[16] Net lending to abroad	38	18	24 (13.1)	40 (15.0)
[17] Final consumption, government	368	385	380 (7.2)	384 (8.0)
[21] Net lending, government	−50	−107	−71 (8.2)	−51 (9.4)
[25] Net lending, corporations	−64	−87	−67 (5.6)	−67 (6.6)
[26] Final consumption, households	1,031	1,102	1,058 (19.6)	1,076 (24.4)
[28] Disposable income, gross, households	1,259	1,435	1,302 (20.4)	1,313 (24.2)
[30] Savings, gross, households	228	333	244 (16.0)	237 (17.6)
[38] Net lending, households	152	211	162 (10.1)	158 (11.1)

Table 4). Prior knowledge of the accuracy of the data is largely ignored at this stage. We have now 40 $(19+16+5)$ pieces of information to estimate the 38 variables, and hence we have two degrees of freedom. In order to identify the 38 variables, two of the exact restrictions are ignored at this stage: supply and use, and other incomes and outlays. The "estimates" obtained in this way do not satisfy these two restrictions, and hence inconsistencies (statistical discrepancies) occur at this first stage.

These statistical discrepancies are eliminated in the second stage of the compilation, by adjusting the estimates of the variables, and indirectly also the values of the indicator ratios used in the first stage. Implicit in the latter procedure are qualitative reliability criteria regarding data and indicator ratios, which the national accountants take into account. In particular the estimates of two variables are adjusted at this stage, which are considered to be less reliable, i.e. gross capital formation and final consumption of households. For a more comprehensive description of the formalized approach to current national accounts compilation practices, the reader is referred to United Nations (2000, V.D.1).

The SNA estimates of the key variables obtained in this way are also presented in Table 7. The discrepancy from our Bayesian estimates is substantial.

5.2. *Key Indicator Ratios*

The key indicator ratios are presented in Table 8. They are closely related to the key variables and include, among others, GDP per capita, household disposable income per capita, GDP real growth, and the share of investment and total

TABLE 8
BAYESIAN ESTIMATES FOR "KEY" INDICATOR RATIOS

	Key Indicator Ratio	Previous Year	Comprehensive Compilation: Current SNA Practice	Comprehensive Compilation: Bayes	Reduced Compilation: Bayes
[28]/[33]	Household disposable income/capita	13,989	15,739	14,305 (300)	14,399 (392)
[4]/[33]	Per capita GDP	20,600	21,944	20,760 (388)	21,246 (544)
[12]/[4]	Export–import gap/GDP (%)	2.2	1.2	1.5 (0.7)	2.2 (0.8)
[3]/[4]	Investment share in GDP (%)	22.3	24.5	22.4 (0.4)	22.4 (0.5)
$([5]-[5]_{-1})/[5]_{-1}$	GDP real growth (%)	***	5.8	0.0 (1.3)	1.5 (3.1)
-[21]/[4]	Government net borrowing/ GDP (%)	2.7	5.3	3.8 (0.4)	2.6 (0.5)
[19]/[4]	Total taxes/GDP (%)	21.8	18.3	21.0 (0.3)	21.8 (0.3)
[26]/[28]	Propensity to consume of households (%)	81.9	76.8	81.2 (1.1)	82.0 (1.3)
[17]/[26]	Government/household consumption ratio (%)	35.7	34.9	35.9 (0.7)	35.7 (0.7)
[26]/[4]	Household consumption/ GDP (%)	55.6	55.1	56.0 (0.7)	55.5 (0.8)
[38]/([38]+[16])	Net lending of households/total net lending (%)	80.0	92.4	87.2 (5.7)	79.7 (5.3)

taxes in GDP. Some of the key indicator ratios ([3]/[4], [26]/[28] and [17]/[26]) are also used as priors, but most are not. The Bayesian estimates of the key indicator ratios are presented under "comprehensive compilation" in Table 8, and can be compared with the SNA estimates. Some of the differences between the SNA and our Bayesian estimates are quite large. For example, we estimate the export–import gap/GDP ratio at 1.5 percent, while the SNA estimate is 1.2 percent. Also, we estimate the government net borrowing/GDP ratio at 3.8 percent, while the SNA estimate is 5.3 percent. The largest difference between the two approaches is in GDP real growth, where our Bayesian estimate indicates zero growth, whereas the SNA approach estimates 5.8 percent growth. The difference is so large, because in current practices, as interpreted here, GDP real growth is solely dependent on the observed growth of output to which it is linked through another indicator ratio, namely the output value-added coefficient. In contrast, the Bayesian approach takes into account the values of all indicator ratios listed in Table 4, and thus also the much more limited growth of population/employment (1.3 percent). This large difference in GDP growth between the two estimation approaches influences, of course, the value of all other indicator ratios that are dependent on GDP. It thus explains the large difference between the two estimates for the export-import gap/GDP ratio (1.5 percent versus 1.2 percent) and net borrowing/GDP (3.8 percent versus 5.3 percent).

In Table 8 we present standard errors of the estimated Bayesian indicator ratios. These standard errors are approximations based on (14). Thus,

$$\text{var}(y/x) \approx \frac{\text{var}(y) + r^2\,\text{var}(x) - 2r\,\text{cov}(y, x)}{\text{var}(x) + (Ex)^2}, \tag{30}$$

where

$$(31) \qquad r = Ey/Ex$$

and Ex, Ey, var(x), var(y), and cov(y, x) denote posterior moments. We see that the export–import gap/GDP ratio and in particular GDP real growth are difficult to estimate precisely.

5.3. *Comprehensive and Reduced Compilation*

Typically, national statistics offices produce not one, but three rounds of estimates of the national accounts variables. Suppose we wish to estimate the variables for year t. The first round takes place in the winter or early spring of year $t+1$. Few actual data are then available, so one has to rely heavily on priors. This round is here called "reduced compilation." The next round may take place one year later and is here called "comprehensive compilation." More data and more accurate data are then available. The final round takes place one year after this.

Our analysis so far can be viewed as the second round ("comprehensive compilation"), where we have access to 19 data. In the first round ("reduced compilation") we have only 8 data: [6], [7], [11], [17], [18], [20], [24], and [32]. The only other difference between the two rounds is that, in the absence of a price survey, the inflation rate is assumed to be 0 percent (rather than 2 percent), so that the GDP price deflator [4]/[5] has a prior mean of 1.5983. We assume that its precision is low (L). The resulting estimates for the key variables and the key indicator ratios are presented in Tables 7 and 8, last panel.

About one half of the key indicator ratios are rather well estimated at the "reduced compilation" stage: household disposable income per capita, investment share in GDP, propensity to consume of households, government/household consumption ratio, and household consumption/GDP. These five indicator ratios are not very sensitive to having more data. Two other indicator ratios (per capita GDP and total taxes as a percentage of GDP) are moderately sensitive, while the other indicator ratios, in particular the GDP real growth rate and the export–import gap/GDP, are poorly estimated at the "reduced compilation" stage. All estimates become more precise when more data are available.

6. Concluding Remarks

The Bayesian estimation approach developed in this paper allows us to obtain estimates for the variables and the indicator ratios between the variables, and also reliability intervals of these estimates. The "simultaneity" feature of this approach introduces several new elements in current national accounts estimation practices. First and most importantly, the approach takes full account of all available information together with the assumed prior precision of that information. Secondly, if we have several pieces of information on the same variable or set of variables, then this causes no problems (multiple priors). Thirdly, all variables and indicator ratios are estimated with their corresponding reliability intervals.

This is new, as variable estimates in national accounts are generally point estimates without estimates of the standard errors.

The basics of the approach have been worked out and illustrated above. Its future potential, however, will depend on the further development of the approach for practical application. Some of these potentials are briefly discussed below.

It is important that through the Bayesian approach a direct link is established between the value and reliability of basic data and the value and reliability of the estimates of the national accounts variables. This link was used above to show how improvements or extensions of basic data sources would lead to reduced reliability intervals of posteriors when additional data sources are employed, and thus to improved reliability of the national accounts estimates. The same link could also be used to quantify the additional reliability that would be achieved by integrating basic data sources through the use of the national accounts framework.

The method also facilitates the simultaneous use of indicator ratios in compilation and analysis. This is important as indicator ratios are generally the core of simple analyses (United Nations, 2000, Chapter 4A). Depending on the availability of basic data at the time or the circumstances that national accounts are compiled, current national accounts estimation approaches use fixed values or point estimates of a few selected indicator ratios as assumptions or "priors" in the national accounts compilation, and "posterior" values of the indicator ratios are close to their "prior" values. This practically eliminates the use of these indicator ratios in analysis. In the Bayesian approach the "prior" and "posterior" values of the selected indicator ratios may be different as the "priors" are not point estimates, but defined with help of reliability intervals. As a result, they can still play a role in analysis, together with other indicator ratios that were not used in the compilation.

The example of the Bayesian approach given in Section 5 is a relatively simple one, based on an aggregated national economic accounts framework with limited scope and detail. This was done solely in order to illustrate the essential features of the method. However, the approach can be applied without any difficulty to a much larger number of variables and indicator ratios and thus to a more realistic accounting framework. There are plans to develop the approach further for use in preliminary accounts when a limited set of basic data is available, to comprehensive annual and benchmark economic accounts, and also to satellite accounts. The further development would be closely aligned with the so-called "systems approach" to macro accounts compilation, which has been developed by the UN Statistics Division and implemented in several countries; see United Nations (1999). An intriguing conclusion from our example is that prior values of indicator ratios may have a larger impact on the posterior estimates of the variables than improved or additional information obtained from new data sources, such as household or enterprise surveys. Of course, this tentative conclusion depends on the assumed precisions of the indicator ratios. It would be interesting to verify the conclusion by further elaboration of the approach for more detailed data systems.

The direct estimation of indicator ratios together with the underlying values of the variables can also be extended to the integration of estimation approaches used in accounting and modeling. In the current approach, indicator ratios are largely defined between variables within the same period. However, the approach could also be applied to accounting covering several periods. This would involve inter-temporal indicator ratios such as growth rates and capital output ratios, which are defined between variables of different periods. Using indicator ratios in this sense would establish a first link between the Bayesian approach and simple modeling based on the use of indicator ratios defined within and between periods. Finally, once the inter-temporal problem has been worked out, the approach could be extended to a more complex link between the Bayesian approach to macro accounts compilation and parameter estimation methods used in econometrics.

7. Appendix: Proof of Theorem 1

The proof proceeds in three steps of increasing generality on the assumed prior distribution. In Lemma A1 we assume that prior information is available on *all* latent variables.

Lemma A1. Let x be an $n \times 1$ vector of latent variables and let d be a $p \times 1$ data vector such that $d|x \sim N_p(Dx, \Sigma)$, where the $p \times n$ matrix D has full row-rank and Σ is positive definite. Suppose that prior information is available in the form $x \sim N_n(q, Q)$, where Q is positive semidefinite (possibly singular). Then the posterior distribution of x is given by $x|d \sim N_n(\mu, V)$, with

$$(32) \quad V = Q - QD'(\Sigma + DQD')^{-1}DQ, \qquad \mu = q - QD'(\Sigma + DQD')^{-1}(Dq - d).$$

Proof. It is obvious and well-known that the posterior of x is normally distributed. Assume first that Q is nonsingular. Then we find the moments μ and V by completing squares:

$$(33) \quad (d - Dx)'\Sigma^{-1}(d - Dx) + (x - q)'Q^{-1}(x - q) = (x - \mu)'V^{-1}(x - \mu) + R,$$

where R does not depend on x. This gives

$$(34) \quad V^{-1} = D'\Sigma^{-1}D + Q^{-1}, \qquad V^{-1}\mu = D'\Sigma^{-1}d + Q^{-1}q,$$

and hence the expressions in the lemma. If Q is singular, the expressions remain valid, because $\Sigma + DQD'$ remains nonsingular. ||

The next step contains the crux of the proof. Here we allow some of the priors to be non-informative (have infinite variance).

Lemma A2. Assume that the conditions of Lemma A1 hold. Assume further that

$$(35) \quad Q = Q_0 + (1/\lambda^2)LL',$$

where Q_0 is positive semidefinite, L has full column-rank ≥ 1, and the identifiability condition $r(DL) = r(L)$ is satisfied. Then, as $\lambda^2 \to 0$, the posterior distribution of x is given by $x|d \sim N_n(\mu, V)$, with

$$(36) \quad V = Q_0 - Q_0 D'\Sigma_0^{-1}DQ_0 + CKC'$$

and

$$\mu = q - (Q_0 + CK)D'\Sigma_0^{-1}(Dq - d), \tag{37}$$

where

$$\Sigma_0 = \Sigma + DQ_0D', \qquad C = I - Q_0D'\Sigma_0^{-1}D, \qquad K = L(L'D'\Sigma_0^{-1}DL)^{-1}L'. \tag{38}$$

Proof. We apply the results of Lemma A1. Letting $R = \Sigma_0^{-1/2}DL$, we have

$$\Sigma + DQD' = \Sigma_0 + (1/\lambda^2)DLL'D' = \Sigma_0^{1/2}(I + (1/\lambda^2)RR')\Sigma_0^{1/2}. \tag{39}$$

The identifiability condition implies that R has full column-rank and hence that $R'R + \lambda^2 I$ has full rank, also at $\lambda^2 = 0$. Now,

$$(\Sigma + DQD')^{-1} = \Sigma_0^{-1/2}(I - R(R'R + \lambda^2 I)^{-1}R')\Sigma_0^{-1/2} \tag{40}$$

and hence

$$\begin{aligned} QD'(\Sigma + DQD')^{-1} &= Q_0D'\Sigma_0^{-1/2}(I - R(R'R + \lambda^2 I)^{-1}R')\Sigma_0^{-1/2} \\ &\quad + L(R'R + \lambda^2 I)^{-1}R'\Sigma_0^{-1/2}. \end{aligned} \tag{41}$$

This gives

$$\begin{aligned} V &= Q - QD'(\Sigma + DQD')^{-1}DQ \\ &= Q_0 - Q_0D'\Sigma_0^{-1}DQ_0 - CL(R'R + \lambda^2 I)^{-1}R'\Sigma_0^{-1/2}DQ_0 \\ &\quad + (1/\lambda^2)CL(I - (R'R + \lambda^2 I)^{-1}R'R)L' \\ &= Q_0 - Q_0D'\Sigma_0^{-1}DQ_0 + CL(R'R + \lambda^2 I)^{-1}L'C' \end{aligned} \tag{42}$$

and

$$\begin{aligned} \mu &= q - QD'(\Sigma + DQD')^{-1}DQ \\ &= q - (Q_0 + L(R'R + \lambda^2 I)^{-1}L' \\ &\quad - Q_0D'\Sigma_0^{-1/2}R(R'R + \lambda^2 I)^{-1}L')D'\Sigma_0^{-1}(Dq - d) \\ &= q - (Q_0 + CL(R'R + \lambda^2 I)^{-1}L')D'\Sigma_0^{-1}(Dq - d). \end{aligned} \tag{43}$$

Letting $\lambda^2 \to 0$ gives the desired results. $\|$

Based on Lemmas A1 and A2 we can now prove Theorem 1.

Proof of Theorem 1. If $m = n$, Theorem 1 follows from Lemma A1 by letting $q = A^{-1}h$ and $Q = A^{-1}HA^{-1\prime}$. If $m < n$, we have less than n "informative" priors. To the m informative priors Ax we now add $n - m$ non-informative priors $L'x$. Since $(A', L)^{-1} = (A^+, L)'$, one verifies that the two statements

$$\begin{pmatrix} Ax \\ L'x \end{pmatrix} \sim N\left[\begin{pmatrix} h \\ 0 \end{pmatrix}, \begin{pmatrix} H & 0 \\ 0 & (1/\lambda^2)I \end{pmatrix}\right] \tag{44}$$

and

$$x \sim N(A^+h, A^+HA^{+\prime} + (1/\lambda^2)LL') \tag{45}$$

are equivalent. Hence, prior information $Ax \sim N(h, H)$ is equivalent to prior information

$$(46) \qquad x \sim N(q, Q_0 + (1/\lambda^2)LL'), \qquad q = A^+h, \qquad Q_0 = A^+HA^{+\prime},$$

when λ^2 approaches 0. Direct application of Lemma A2 now yields the required results. $\|$

References

Barker, T., F. van der Ploeg, and M. Weale, A Balanced System of National Accounts for the United Kingdom, *The Review of Income and Wealth*, 30, 461–85, 1984.

Byron, R. P., The Estimation of Large Social Account Matrices, *Journal of The Royal Statistical Society, Series A*, 141, 359–67, 1978.

Clemen, R. T. and R. L. Winkler, Combining Probability Distributions from Experts in Risk Analysis, *Risk Analysis*, 19, 187–203, 1999.

Genest, C. and J.V. Zidek, Combining Probability Distributions: A Critique and Annotated Bibliography, *Statistical Science*, 1, 114–48, 1986.

Goldstein, M., Adjusting Belief Structures, *Journal of the Royal Statistical Society, Series B*, 50, 133–54, 1988.

Magnus, J. R. and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, revised edition, John Wiley, Chichester/New York, 1999.

Morgenstern, O., *On the Accuracy of Economic Observations*, second edition, Princeton University Press, Princeton, 1963.

O'Hagan, A., *Bayesian Inference*, Kendall's Advanced Theory of Statistics, Volume 2B, Edward Arnold, London, 1994.

Rao, C. R., *Linear Statistical Inference and Its Applications*, 2nd edition, John Wiley, New York, 1973.

Stone, R., D. G. Champernowne, and J. E. Meade, The Precision of National Income Estimates, *Review of Economic Studies*, 9, 111–25, 1942.

United Nations Development Program, *Human Development Report*, Oxford University Press, New York/Oxford, 1996.

United Nations, Eurostat, IMF, OECD, and World Bank, *A System of National Accounts*, New York, Brussels/Luxembourg, Paris, Washington, DC, 1993.

United Nations, Statistical Office, *A System of National Accounts*, Studies in Methods, Series F, Volume 2, No. 3, United Nations, New York, 1968.

United Nations, Statistics Division, *Handbook of National Accounting: A Systems Approach to National Accounts Compilation*, Studies in Methods, Series F, No. 77, United Nations, New York, 1999.

———, *Handbook of National Accounting: Use of Macro Accounts in Policy Analysis*, Studies Methods, United Nations, New York, 2000, forthcoming.

Van der Ploeg, F., Reliability and the Adjustment of Sequences of Large Economic Accounting Matrices (with discussion), *Journal of the Royal Statistical Society, Series A*, 145, 169–94, 1982.

———, Economic Policy Rules for Risk-Sensitive Decision Making, *Zeitschrift für Nationalökonomie*, 44, 207–35, 1984.

———, FIML Estimation of Dynamic Econometric Systems from Inconsistent Data, *International Journal of Systems Science*, 16, 1–29, 1985.

Van Tongeren, J. W., Development of an Algorithm for the Compilation of National Accounts and Related Systems of Statistics, *The Review of Income and Wealth*, 31, 25–47, 1985.

Wiper, M. P. and S. French, Combining Experts' Opinions Using a Normal-Wishart Model, *Journal of Forecasting*, 14, 25–34, 1995.

World Bank, *World Development Report*, *Investing in Health*, Oxford University Press, New York/Oxford, 1993.

Chapter 6

Bayesian integration of large SNA data frameworks with an application to Guatemala

Bayesian integration of large SNA data frameworks with an application to Guatemala*

Jan W. van Tongeren
Former Head of the National Accounts Branch,
United Nations Statistics Division, New York

Jan R. Magnus
Department of Econometrics & Operations Research,
Tilburg University

Abstract: We present a Bayesian estimation method applied to an extended set of national accounts data and estimates of approximately 2500 variables. The method is based on conventional national accounts frameworks as compiled by countries in Central America, in particular Guatemala, and on concepts that are defined in the international standards of the System of National Accounts. Identities between the variables are exactly satisfied by the estimates. The method uses ratios between the variables as Bayesian conditions, and introduces prior reliabilities of values of basic data and ratios as criteria to adjust these values in order to satisfy the conditions. The paper not only presents estimates and precisions, but also discusses alternative conditions and reliabilities, in order to test the impact of framework assumptions and carry out sensitivity analyses. These tests involve, among others, the impact on Bayesian estimates of limited annual availability of data, of very low reliabilities (close to non-availability) of price indices, of limited availability of important administrative and survey data, and also the impact of aggregation of the basic data. We introduce the concept of 'tentative' estimates that are close to conventional national accounts estimates, in order to establish a close link between the Bayesian estimation approach and conventional national accounting.

*We are grateful to Magda Ascues, Carlos Blanco, Victor Flores, Arthur Giesberts, and Ruud Picavet for helpful comments and preparation of documents. Responsibility for the content of this paper is with the authors and does not reflect in any way the point of view of the UN Statistics Division.

1 Introduction

This paper describes the use of a Bayesian estimation approach in the compilation of national accounts. The application is based on a project carried out in Central America and the results are presented for one country, namely Guatemala. The compilation involves approximately 2500 variables, which is close to what is conventionally involved in an extended compilation based on the international standards of the System of National Accounts (SNA). The basis of the SNA compilation was mainly the 1993 version of the SNA; the 2008 version was not fully implemented in the countries that participated in the project; see United Nations *et al.* (1993, 2008).

The compilation of SNA data with the help of a Bayesian estimation approach builds on what was developed in Magnus *et al.* (2000), Magnus and Van Tongeren (2002), and Danilov and Magnus (2008). The paper presents, for the first time, a real-life application to a large and realistic data set. Moving from a small data set (40 variables in our 2000 paper) to a large data set (2500 variables) is a big step. The experiences acquired in the application of the Bayesian approach to several countries in Central America led to many improvements in the method and the software. In particular, extensive use is now made of 'sparse' matrix theory in the SNAER (System of National Accounts Estimation and Reconciliation) software in order to increase the accuracy and speed of the estimation process.

The Bayesian approach is applied to 'frameworks' of data and estimates, which can be used both in compilation and analysis. The frameworks are matrices or groups of matrices, in which two types of relations are defined between the variables: ratios and identities. The identities are based on SNA definitions and SNA balances; the ratios between variables are similar to the compilation ratios used by national accountants and also reflect simple ratio analyses carried out by analysts using the national accounts and other data. The frameworks make explicit the prior reliabilities of available data and ratio values that reflect the degree of willingness to change values, based on implicitly perceived trust in these values.

The BSNA (Bayesian SNA) framework used in the present study was developed for six Central-American countries, as part of a project sponsored by the Netherlands Organization for International Cooperation in Higher Education (NUFFIC) in cooperation with the Instituut voor Ontwikkelingsvraagstukken at Tilburg University, The Netherlands, and the Consejo Monetario Centroamericano (CMCA). The countries consist of the six member states of CMCA: Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, and the Dominican Republic. The data in this paper refer to the Guatemala BSNA for 2005 as the benchmark year, and 2006 as the current year.

The objective of the Bayesian method is the same as in conventional national accounting, where it is called ‘reconciliation’ or ‘integration’ of data. But Bayesian estimation or integration differs from the conventional national accounting approaches in several respects. First, all conditions of conventional estimates are formalized: identities are explicitly included and ratios are introduced as priors; second, reliabilities of data and ratio values are reflected in well-defined prior variation coefficients; third, the system is simultaneous rather than sequential; and finally, updating the system when new information becomes available is easy and fast, and does not require changes in the compilation method.

In the estimation process priors and data are combined to a posterior distribution. The mean of the posterior distribution is then taken as the estimate and the variance in the posterior distribution as a measure of precision. In this way, Bayesian estimates of all variables of the framework are derived, also for the variables for which no basic data are available and also for the ratios. In addition, standard deviations of all estimates and ratios defined in the framework are obtained.

The plan of this paper is as follows. In Section 2 we describe some characteristics of the economy of Guatemala, and discuss data availability and how conventional national accounts are compiled. These economic characteristics and the data compilation approaches are reflected in the design (or ‘architecture’; see Jorgenson, 2009; Jorgenson *et al.*, 2010; and Vanoli, 2010) of the BSNA data framework, presented in Section 3. Particular attention is paid to the use of ISIC and CPC classifications in the BSNA framework and to the SNA transactions that are incorporated in the framework and the types of analyses that they support. Section 4 provides details on the Bayesian inputs that are used in the compilation: ratios, identities, and reliabilities of basic data and ratio values. In Section 5 we summarize the Bayesian estimation method and how this method is reflected in two software programmes, SNAER and INTERFACE, that are used to arrive at Bayesian estimates. Section 6 presents the results of a number of tests of the framework, including tests that measure the impact on the Bayesian estimates if fewer data are annually available, and tests of different prior reliabilities assigned to basic data and ratio values. This section also describes the use of so-called 'tentative' estimates that are close to conventional national accounts estimates and that result in improved Bayesian estimates. Section 7 describes the results of a number of sensitivity tests. These tests quantify the impact on the Bayesian estimates of different scopes of the framework, of different availability of basic survey and administrative data, and of aggregation of basic data. They also quantify the impact of these alternative options on the posterior reliability of the Bayesian estimates. Section 8 summarizes our experiences with the Bayesian estimation method, and provides some suggestions for further work.

2 Characteristics of Guatemala

Guatemala is a small country in Central America, about 5% of the size of Mexico. It is bordered by Mexico to the north and west, the Pacific Ocean to the southwest, Belize to the northeast, the Caribbean to the east, and Honduras and El Salvador to the southeast. With a population of close to 14 million, it is the most populous of the Central-American countries. Its per capita GDP ($2115 in 2005) is about one quarter of Mexico's. Coffee, sugar, and bananas are the main products. The Central American Free Trade Agreement between Guatemala and the USA, signed in 2006, has increased investment in the export sector. Guatemala has an important 'free trade' zone, in which semi-manufactured products are temporarily imported, processed, and then exported again. Guatemala's economy is dominated by the private sector, which generates about 85% of GDP. Most manufacturing is light assembly and food processing, geared to the domestic, US, and Central-American markets. In addition to the main exports (coffee, sugar, bananas), the exports of textiles, apparel, and non-traditional agricultural products (mainly winter vegetables, fruit, and cut flowers) have boomed in the past few years, as of course has tourism. The revenues of migrant workers, employed mainly in the USA, are an important element of household disposable income. The distribution of income remains highly unequal; in fact, the most unequal in the region. The government sector is small with its business activities limited to public utilities, ports and airports, and several development-oriented financial institutions. The economy's structure is reflected in the International Standard Industrial Classification of All Economic Activities (ISIC) and Central Product Classification (CPC) breakdowns of the BSNA framework, as explained in Section 3.

As in other countries in Central America, national accounts in Guatemala are compiled by the Central Bank. The main focus of the national accounts compilation is the Supply and Use Table (SUT), but it also includes the compilation of Integrated Economic Accounts (IEA). The Central Bank also compiles financial accounts and balance sheets for monetary analysis. The two compilation activities are carried out by different departments in the Central Bank and managed by staff with different specializations: national accountants versus monetary economists and modelers. Communication between the two types of specialists is often difficult.

Since an extension of BSNA to monetary accounts in the future is likely, the BSNA framework pays special attention to the development of the IEA and to sectorization therein, emphasizing in particular the three pillars of monetary-fiscal-financial analysis: the government sector, the financial corporate sector, and the rest of the world. The remaining sectors are further broken down for national accounts purposes into private and public non-financial corporations, households, and non-profit institutions. By making households a separate sector in the IEA, the framework distinguishes, within the ISIC classification by industries, between large production units that belong to the non-financial corporation sector and small production units that belong to the household sector. It is then also possible to distinguish between a measure of net product, i.e. GDP generated by the economy, and household disposable income, which is that part of GDP that reaches households and thus can be considered as an approximate measure of well-being. In addition, through concepts of household saving and net lending, we can then measure the contribution of households to the financing of investments.

3 The Bayesian SNA framework

The BSNA framework is an SNA framework designed in EXCEL format. The EXCEL cells do not only represent values of variables in the framework, but also identities and ratios defined between the variables, and reliabilities attached to the values of the variables and ratios. The degrees of freedom in designing the BSNA framework were not large, because the data to which the Bayesian method has been applied are based on what the 1993 SNA recommends in terms of transaction and transactor concepts. The framework is mainly redesigned in terms of the classifications used and the variables included, representing specific features of Guatemala's economy.

FIGURE 1

A synopsis of the framework is presented in Figure 1. Cells with different types of information are distinguished by different shades. There are four types of information in the framework, namely:

- Variables for which basic data are available;
- Variables for which no basic data are available, but have values that are derived, using base year structures and SNA identities;
- Ratios that are defined between the variables of the framework, i.e. variables with or without basic data. Both ratio definitions and values are used; and

- Identities that are defined between the variables of the framework. These are defined in such a manner that, if satisfied, identity cells have value = 0. Both identity definitions and identity values are used.

BSNA Guatemala is represented by six alternative frameworks for 2005 and 2006, which all have the same format as Figure 1. These are referred to as alternative options in Table 3 and in the tests and sensitivity analyses of Sections 6 and 7.

2005GV is the framework for the benchmark year 2005. It includes a complete data set with all values of variables treated as basic data (colored green (G) in the EXCEL frameworks). The values (V) of all framework cells are known. The benchmark framework does not only include basic data, but also estimates that are made by national accountants on the basis of similar assumptions (ratios and identities) as are used in the BSNA compilation. All cells in the benchmark framework are treated as basic data for BSNA purposes.

2005YF simulates for 2005 limited data availability in annual estimates. Only those values of variables are treated as basic data that are normally available annually, when using administrative data sources and surveys. The values of the remaining variables are treated as not available (framework cells are colored yellow (Y), meaning that those values are only dealt with to a limited extent in the Bayesian estimation). The latter variables have so-called tentative values, as explained in Section 6.3. They are derived with the help of a subset of formulas (F) of ratios and identities and 2005 ratio values, and are therefore the same as the corresponding values in the 2005GV framework.

2006GV includes a nearly complete data set for 2006, which is based on 'estimates' of the conventional SNA compilation. Nearly all framework cells are treated as basic data, except some that are not covered in the SNA conventional compilation. Values for these variables are derived indirectly, using 2005 data structures, and are not treated as basic data. This applies in particular to the breakdown for intermediate and final uses between local output and imported products.

2006YF is similar to the 2005YF framework, except that 2006 and not 2005 data are used as basic data. The values of the remaining variables without basic data are derived from the 2006 basic data with the help of the same ratios and identities that were used in the 2005YF framework. Identities not used in the latter derivation may not be equal to zero, and ratios that are not used may have values that are different from the 2005 values.

2006GF uses the same formulas as in the 2006YF framework, but instead of treating the derived values as variables without basic data, they are treated as basic data (colored green), and are therefore also assigned prior reliabilities.

2006YV(2005) refers to the framework with basic data for 2006. If no basic data are available in 2006, we use 2005 data.

The coding of the framework names (GV, YF, YV, GF) is based on the color codes used in the INTERFACE programme, discussed in Section 5. Green (G) is used for variables with basic data and yellow (Y) for variables without basic data. When actual values are used this is indicated with V, and when ratio or identity formulas are used to derive the values of cells for which no data are available this is indicated with F.

All frameworks have identical formats, as presented in Figure 1, in terms of the columns and rows for which transactor and transaction categories are presented. Each includes the two main data segments of the SNA, i.e. the Supply and Use Table (SUT) in current and constant prices and the Integrated Economic Accounts (IEA). In Figure 1 the SUT is presented on the left-hand side and the IEA on the right-hand side. Both tables are matrices, cross-classifying ‘transactor’ categories in the columns with ‘transaction’ categories in the rows. Incorporation of IEA is emphasized in the BSNA framework, in order to facilitate a link with monetary analysis. Also included in the framework is the cross-classification between industries and sectors (CCIS), linking the SUT and IEA; it is presented below in Figure 1 in the middle of the SUT segment.

Figure 1 only includes aggregate sectors, but in the extended BSNA framework used in the compilation, nine SNA sectors are distinguished:

- government (GOV);
- financial corporations (FC): Central Bank, deposit money banks, and other;
- non-financial corporations (NFC): public (NFCpu) and private (NFCpri);
- non-profit institutions serving households (NPI);
- households (HH); and
- rest of the world (ROW).

For the SUT the columns refer to a combination of industry (ISIC) and product (CPC) groupings. The format of the SUT differs from the 1993 SNA; see Van Tongeren (2004). In this adapted SUT format, the supply and use product rows are transposed to column format and combined with the industry columns of the 1993 SNA-SUT format. Thus, supply and use data by CPC categories of products (imports and exports and final consumption and capital formation) and industry accounts by ISIC categories (output, intermediate consumption by use, value added, value added components, and employment) in current and constant prices are combined in the same column. This combination of CPC and ISIC categories in the columns is indicated by ISIC and/or CPC to the left of the SUT in Figure 1. If ISIC is indicated, the row details concern industries, and if reference is made to the CPC, row details concern products. This distinction between industry and product categories is particularly relevant for output, which is presented by industries and by products. If one industry only produces one product, industry and product categories are interchangeable, and output in terms of industries and products is identical. But this one-to-one correspondence does not hold if industries produce secondary products that are characteristic of other industries. The secondary products are presented in Figure 1 between rows of output by industry and product, where they are classified by type of products in the rows and by industries producing those products in the columns. The totals of output by products in current and constant prices for each industry/product column are then derived by adding to industry output corresponding products produced by other industries in other columns and deducting secondary products produced by the industry in the column.

International standards on ISIC and CPC have been used in the framework. As a minimum two-digit breakdowns of ISIC and CPC have been incorporated, following common practice. A number of special features are reflected in the classification. When designing the country-specific classifications, there is close coordination between the details of ISIC industry and CPC-related product categories and also between sector and industry categories, so that each product category is assigned to a unique ISIC category and each ISIC category is assigned to a unique sector category. In the classifications thus designed, the one- to four-digit levels of the classifications may be combined, but the two-digit level is always maintained in support of international comparisons. Another feature is the introduction, within the product-related CPC classifications, of product distinctions that identify origins (output and imports) and destinations of products (intermediate consumption, final consumption, and capital formation). Introducing these product distinctions, which are based partly on the UNSD (1989) Classification by Broad Economic Categories considerably simplifies the supply-use identities and thus facilitate the reconciliation of data in which these identities play a role.

In some cases, special features of the SUT and IEA segments of the framework are introduced that differ from those of the conventional SNA. We mention four such deviations. First, within the destination of products in intermediate consumption, final consumption, and capital formation, a distinction is made between products originating from local output and imports, so that the import dependence of the economy can be measured in detail. Second, the framework includes many details on secondary output, which facilitate measuring the link between output by industry and product. Third, a broad set of price indices by product categories is included in the framework, linking the current and constant price product flows. Finally, the framework includes a link between GDP and household disposable income, as a means of refocusing aggregate analysis from sole dependence on the GDP production measure to a balanced focus which also includes an SNA measure of well-being. The latter is in line with the recommendations in Stiglitz *et al.* (2009). Details of these special classification and SNA transaction features are discussed in Van Tongeren (2010).

4 Bayesian conditions in the BSNA framework

4.1 Bayesian conditions

The Bayesian conditions include four elements: basic data, ratios, identities, and the coefficients of variation of the basic data and ratio values. Ratios, ratio values, and identities are used in the benchmark compilation to ensure that the Bayesian estimates are compatible. In the annual compilation they are used to supplement limited data availability, thus facilitating the estimation of variables for which no basic data are available. The scope of ratios is the same in the benchmark and the annual compilation. The SUT ratios on the left-hand side of Figure 1 include price indices, input-output coefficients, use coefficients, wage and mixed income rates, and value added and other industry distribution coefficients. The IEA ratios on the right-hand side include coefficients describing the composition of household disposable income, the finance of capital formation, and the distribution between the sectors of the IEA revenues and expenditures. The SUT identities on the left-hand side of Figure 1 refer, among others, to CCIS-IEA identities for the production accounts, supply-use identities for products, identities defining value added and operating surplus, the identity between output of trade and transport, and the sum of trade and transport margins, as well as overall GDP identities. The identities in the IEA segment of the scheme include the identities defining disposable income, saving, and net lending of sectors. To the right of the IEA are included identities between totals and details of SUT rows, identities between revenues and expenditures in the IEA, and identities between variables of IEA and SUT.

In the annual scheme 2006YF there are 2719 variables, supported by 531 basic data, 1120 identities, and 2294 ratios. Thus, while only 19.5% of the variables are supported by basic data, per variable 1.45 information items are available. Such a large number of information items per variable allows checks to be made between basic data and priors, which is close to how conventional national accounts practices are carried out.

The difference with conventional accounting practices concerns the number of ratios and identities that can be taken into account, and the use of reliabilities. In conventional national accounting the number of assumptions is generally equal to the number of variables without basic data, while in the Bayesian estimation approach of SNAER any number of restrictions could be accommodated. (This feature of conventional national accounting is used in the derivation of so-called 'tentative' estimates prior to Bayesian estimation in the present method; see Section 6.3.) The large number of restrictions in the Bayesian estimation approach serves as a means of checking the restrictions. This is a feature that is not explicitly available in conventional national accounting practices. Reliabilities are used implicitly when national accountants adjust data, but they are explicitly used in the Bayesian estimation, by adjusting prior values of variables and ratios more when their reliability is low (large standard deviation) and less when their reliability is high (small standard deviation). If the number of basic data items increases, which will happen between preliminary and final estimation, the number of information items will grow and more checks will become available, while the number of ratios and identities will remain the same. In the ideal case, there will be basic data for all cells, and this is simulated by the 2005GV and 2006GV frameworks; see the first test in Section 6.2. In the case of the 2006GV framework, there are 2550 basic data, 1167 identities, and 2422 ratios for 3037 variables, which means that for each variable 2.02 information items are available for Bayesian estimation and checks. In the next sections the scope of basic data, ratios, and identities is reviewed in more detail, and also attention is paid to the reliabilities attached to those values. The focus is on the current estimates for the annual scheme 2006YF.

4.2 Variables with and without basic data

In the annual frameworks, eight types of basic data are assumed to be available.

TABLE 1

These are presented in Table 1, which also includes annotations as to how they have been compiled, and indicates the prior reliability of each. This reliability, to be determined by the national accountant, is expressed as a variation coefficient, which is the inverse of the t-ratio. The table shows that the data that are considered to be most reliable (F, nearly fixed) are all data on output of goods and services based on economic surveys, the totals of exports and imports, total employment, the administrative data of the government, financial corporate, and rest of the world sectors, and also price indices. High (H) reliability is accorded to the total of household final consumption based on household surveys, as well as the detail of exports and imports based on foreign trade and Balance of Payments Statistics. Medium (M) reliability is accorded to the detail of household final consumption by products and the detail of employment by economic activities. In the benchmark and other schemes with more basic data, other prior reliabilities are assigned to other data. For example, superior (S) reliability is assigned to sector data of public non-financial corporations, and financial corporations other than Central Bank and deposit money banks, and low (L) reliability is assigned to production account items other than output, and also to trade and transport margins on products. Poor (P) reliability is not assigned to any item. All data in the benchmark scheme for the household, non-profit institution, and private non-financial corporation sectors are assigned low (L) reliability.

4.3 Identities and ratios

In addition to the basic data, the BSNA includes definitions of identities, and definitions and values of ratios.

TABLE 2

Their location and specific functions are described in Table 2. The scope of identities and ratios does not differ between benchmark and annual schemes. Also included in the table are indicators of the prior reliabilities of ratio values. The highest reliability (superior, S) has been assigned to input-output coefficients, industry-sector distribution coefficients, coefficients of household final consumption distribution by products, and coefficients of distribution of value added by industry, while the lowest (L) reliability has been assigned to IEA coefficients of distribution of revenues and expenditures between sectors. No prior reliabilities have been assigned to aggregate ratios, i.e. GDP growth rate, propensity to consume, household disposable income/GDP ratio, and the terms of trade (PX/PM). The last column of the table identifies the ratios and identities that are used in the derivation of values for variables for which no basic data are available. They are referred to as 'tentative' estimates in Section 6.3, where they will be explained further.

The incorporation of ratios (and also identities) is much determined by the availability of data and the scope and design of the framework based thereon. This is clearly shown by comparing the design of our BSNA framework with the architecture of the US framework as presented in Jorgenson (2009). The latter framework includes data on the stocks of fixed assets and thus allows for the separate incorporation of rates of return and capital productivity ratios, which are not included in the present scheme. By incorporating constant price data up to saving it is possible to include not only growth rates of output (GDP), but also measures of the increase in the level of well-being.

5 Bayesian method and software

5.1 The Bayesian method

In contrast to the classical (frequentist) approach, a Bayesian does not assume 'true' parameters (latent variables) x. Instead, a probability distribution of the parameters is assumed, the so-called prior distribution. The data then serve to modify the prior idea of the 'truth' into a more complete idea: the posterior distribution. The mean of the posterior distribution can then be viewed as an 'estimator' of x, and the variance of the posterior distribution serves as a measure of its precision. When both the likelihood and the priors are based on the normal distribution, the posterior is normal as well, and therefore there is no mathematical difference between data and priors, although there is of course a conceptual difference. This simple observation leads to equivalences which are utilized in our software.

In this section we summary the mathematics underlying our approach, and describe the two computer programs, SNAER (System of National Accounts Estimation and Reconciliation) and INTERFACE, which together provide estimates and precisions of the latent variables. Our problem is complex because we encounter matrices (data and restrictions) that are large and sparse. A matrix is 'sparse' when it has many structural zeros, and it is 'large' when we have, say, 2^{11} variables and 2^{13} observations, thus giving $2^{24} \approx 16.8$ million entries in the design matrix.

5.2 Data, priors, and linear restrictions

In the formal statistical framework we consider a vector x consisting of n latent variables $x_1, x_2, \ldots, x_n$. Data are available on p components (or linear combinations) of x. Let d_1 denote the $p \times 1$ data vector. Our starting point is a measurement equation,

$$d_1|x \sim \mathrm{N}_p(D_1 x, \Sigma_1), \tag{1}$$

which tells us that the conditional distribution of d_1 given x is normal with a mean which is linear in x and a variance which does not depend on x. Typically, the $p \times n$ matrix D_1 is a selection matrix, say $D_1 = (I_p, 0)$, so that $D_1 x$ is a subvector of x, but this is not required. Neither is it required that the matrix D_1 has full row-rank. Measurements are unbiased in the sense that $\mathrm{E}(d_1|x) = D_1 x$. The $p \times p$ matrix Σ_1 denotes a positive definite variance matrix, typically (but not necessarily) diagonal.

In addition to the p data, we have access to two further pieces of information: prior views concerning the latent variables or linear combinations thereof, and deterministic linear constraints. More specifically, we have m_1 random priors:

$$A_1 x \sim \mathrm{N}_{m_1}(h_1, H_1) \tag{2}$$

and m_2 exact restrictions (identities):

$$A_2 x = h_2 \text{ (almost surely)}, \tag{3}$$

in total $m = m_1 + m_2$ pieces of prior information.

We assume that the $m_1 \times m_1$ matrix H_1 is positive definite (hence nonsingular) and that the $m_2 \times n$ matrix A_2 has full row-rank m_2 (so that the exact restrictions are linearly independent and thus form a consistent set of equations). We define

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \quad H = \begin{pmatrix} H_1 & 0 \\ 0 & 0 \end{pmatrix},$$

and assume that $\mathrm{rk}(A) = m$, which implies of course that both A_1 and A_2 have full row-rank. The rank condition on A is not a serious restriction, because we can freely move priors to data (and vice versa). Hence the condition $m \leq n$ is not restrictive either.

In order to identify all n variables from the information (data and priors) we need at least n pieces of information: $m+p \geq n$. But this is not sufficient for identification, because some of the information may be on the same variables.

Necessary and sufficient for identification is the condition

$$\operatorname{rk}\begin{pmatrix} A \\ D_1 \end{pmatrix} = n,$$

which is automatically satisfied when $m = n$.

5.3 Estimation: the SNAER software

There are several equivalent ways to estimate the components of x and their variances (Danilov and Magnus, 2007). The equivalence is based on two facts. First, a Bayesian analysis with normal data and normal priors is closely linked with a quadratic minimization problem. Second, best linear unbiased estimation is closely linked to quadratic minimization (least squares). A Bayesian solution is provided in Theorem 1 of Magnus *et al.* (2000), but it involves Moore-Penrose inverses and is not easily computable for large sparse systems. An easier, but equivalent, solution is obtained by using the close relationship between best linear unbiased estimation and least squares (Rao, 1971, 1973). Defining

$$D = \begin{pmatrix} D_1 \\ A_1 \end{pmatrix}, \quad d = \begin{pmatrix} d_1 \\ h_1 \end{pmatrix}, \quad H = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & H_1 \end{pmatrix},$$

we obtain estimates of x by solving the constrained problem

$$\begin{aligned} &\text{minimize } (d - Dx)'\Sigma^{-1}(d - Dx) \\ &\text{subject to } A_2 x = h_2. \end{aligned} \tag{4}$$

This can be simplified by writing

$$A_2 = (A_{21} : A_{22}),$$

where A_{21} is an $m_2 \times (n - m_2)$ matrix and A_{22} is a *non-singular* $m_2 \times m_2$ matrix. Partitioning x correspondingly, we can write the restriction as

$$A_{21}x_1 + A_{22}x_2 = h_2,$$

so that

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} I_{n-m_2} \\ -A_{22}^{-1}A_{21} \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ A_{22}^{-1}h_2 \end{pmatrix} = Qx_1 + q.$$

Let

$$B = \Sigma^{-1/2}D = \begin{pmatrix} \Sigma_1^{-1/2}D_1 \\ H_1^{-1/2}A_1 \end{pmatrix}, \quad b = \Sigma^{-1/2}d = \begin{pmatrix} \Sigma_1^{-1/2}d_1 \\ H_1^{-1/2}h_1 \end{pmatrix}.$$

Then the constrained problem (4) can be written as the unconstrained problem

$$\min_{x_1} \ \|(b - Bq) - BQx_1\|^2,$$

which in turn can be rewritten as

$$\min_{x_1} \left(\frac{1}{2}x_1'\Gamma x_1 - c'x_1 \right), \tag{5}$$

where $\Gamma = Q'B'BQ$ and $c = Q'B'(b - Bq)$. This is the format in the SNAER software, which relies on Harwell's HSL_VF06 procedure which, in turn, is based on the Gould-Nocedal (1998) algorithm.

The Harwell routine does not provide the variance matrix, elements of which can be computed as follows. The j-th column $v^{(j)}$ of the matrix $V = \Gamma^{-1}$ can be found by minimizing $\|\Gamma^{-1/2}e^{(j)} - \Gamma^{1/2}v^{(j)}\|^2$ for all j, that is, by solving the quadratic minimization problem

$$\min_{v^{(j)}} \left(\frac{1}{2}v^{(j)\prime}\Gamma v^{(j)} - e^{(j)\prime}v^{(j)} \right) \qquad (j = 1, \ldots, n - m_2),$$

using the Gould-Nocedal algorithm, where $e^{(j)}$ denotes a vector all whose components are zero except the j-th component which is one. The value at the minimum equals $-v_j^{(j)}/2$. This method is specifically designed for sparse systems.

Thus we solve the constrained problem (4) in two steps. First we identify and invert a non-singular submatrix A_{22} from A_2; then we solve (5). The two-step procedure has the advantage that the dimension of the system is much reduced. Moreover, the first step can be done once and then the results of the reduction may be used for many restricted least-squares problems. In particular, the A_2 matrix is usually fixed because it represents the structure of the economy, while the matrices related to A_1 are priors that will vary.

In our case, some of the priors are so-called indicator ratios. These priors are non-linear and hence need to be linearized. Suppose we have a prior indicator ratio $R = y/x \sim (r, \tau^2)$. We wish to replace this non-linear prior by its linearization $y - rx \sim (0, \omega^2)$. The question is how to choose ω^2. This question is discussed in Danilov and Magnus (2008, Section 6), where an invariant linearization method is proposed. Invariance here means that we obtain the same prior whether we start from y/x or from x/y.

Summarizing, the information in our system of latent variables to be estimated consists of incomplete data (with precisions), priors on a subset of the variables or linear combinations thereof (with precisions), and exact linear restrictions. Our system is large and sparse. It is large, because we may have 2^{11} variables and 2^{13} observations, thus giving $2^{24} \approx 16.8$ million entries in the design matrix. It is sparse, because information is often available on one variable at a time, and restrictions are often definitions involving only a small number of variables. The SNAER software, especially developed for this project and described in some detail in Danilov and Magnus (2008), provides estimates and variances (and desired covariances) of the complete x-vector. The method takes full account of all accounting identities, the solutions are continuous rather than discrete, and multiple priors on variables or linear combinations of variables are allowed. The posterior estimates take all prior and data input into account and come with precisions. The system is transparent, flexible, and fast.

5.4 The INTERFACE

The operational aspects and the mathematics of the approach are reflected in two computer programmes: SNAER and INTERFACE. The SNAER program was discussed in the previous subsection. It requires inputs (data, prior indicator ratios, identities, and prior reliabilities) in a specific format. The INTERFACE converts the inputs of the EXCEL worksheets into a format that is accepted by SNAER. In the process of this conversion, INTERFACE carries out a number of checks on the data and Bayesian inputs, before inputting this information into SNAER for further processing.

The INTERFACE is written for EXCEL spreadsheets. It extracts information of any framework that is defined in EXCEL. Thus, it reads and extracts the values of basic data, and of ratio formulas and values included in the framework files. The formulas in those cells are recorded by INTERFACE and converted to a format that can be read in SNAER. It reads and extracts separately the values of the ratio cells. If those values are not applicable, it reads and extracts instead (with preference) ratio value information that is included in EXCEL comments attached to the (purple-colored) cells. Furthermore, it reads and extracts the formulas of identities that are identified in the EXCEL framework with the help of (blue-colored) cells, and also converts those formulas to a format that can be read in SNAER. Finally, it reads and extracts information on reliabilities of basic data and values of ratios.

The reliabilities are expressed as ‘fixed’ (F), ‘superior’ (S), ‘high’ (H), ‘medium’ (M), ‘low’ (L), or ‘poor’ (P), and they are expressed in the framework as percentages of coefficients of variation. Thus, if the reliability of a basic data item or ratio value in the BSNA framework is poor, it may deviate in the final Bayesian estimates 24% from the original value. Similarly, it may deviate 12% if the reliability is 'low', 6% for medium reliability, 3% for high reliability, and 1% for superior reliability.

In the process of extracting the information for use in SNAER, the INTERFACE also carries out data checks and presents error messages if necessary. The three most important checks are the following. First, identities should include at least one variable of the framework. (This limitation to one variable is included for practical purposes, as many of the identities are defined as EXCEL sums, in which many cells are not variables in the system.) Second, both the numerator and the denominator of ratios should be available as variables in the framework. Third, all basic data and ratio values must have been assigned reliabilities.

After all error messages have been cleared, the INTERFACE transfers data to the SNAER input files. After SNAER has calculated estimates and precisions, these are converted back into the format of the framework, so that the user can assess the resulting estimates in their framework context.

6 Tests of framework

6.1 Comparison of framework options

We now present various Bayesian estimates for the years 2005 and 2006. The estimates—referred to as options in Table 3—enable us to assess the impact of changing the conditions embedded in the framework on the Bayesian estimates. These impacts as analyzed below and in Section 7.

TABLE 3

Table 3 summarizes the results of twenty-one different options. The options are described briefly, highlighting deviations from the six standard frameworks described in Section 3. The quantitative assessment of the estimates in subsequent tables is done in terms of major aggregates of GDP and household disposable income, even though the underlying materials would permit making the assessments for all details of variables and ratios in the BSNA framework. There are seven columns in Table 3. The base year is 2005 and the current year is 2006.

In column 2 we highlight the deviations from one of the four standard frameworks: options 1, and 12–14. The standard format for 2005 (option 1) and 2006 (option 13) includes data for all cells, which are treated as basic data. These include changes in inventories and a full set of secondary products. Standard prior precisions are applied, as presented in Tables 2 and 3. The 2005 and 2006 'data' used in the two frameworks are different. The 2005 benchmark data are based on the conventional practice of using available data in a detailed compilation process in order to establish new data structures for use in future compilations. The 2006 full data set is also based on conventional practices, but applying a more limited data analysis, and using mainly data structures of 2005 to estimate the values of the remaining variables and ratios. The standard format of the annual 2005 or 2006 frameworks (options 12 and 14) includes only a limited data set that is annually available on output, employment, imports and exports, price indices, and data on the government, financial corporate, and rest of the world sectors. The annual frameworks do not include changes in inventories and include only a limited set of data on secondary products. Assumptions based on selected ratios and identities are used to arrive at tentative estimates for missing data. Only annually available data are treated as basic data. Values derived on the basis of assumptions are not considered available and therefore not treated as basic data (except in the 2006GF framework of option 15); they are, however, used in the linear approximation of ratios. Standard precisions are applied to basic data and ratio values.

Columns 3 and 4 present the number of 'distortions' in the Bayesian estimates. These are assumed to occur in two instances. First, when Bayesian estimates differ more than 2% in absolute terms (column 3) from the conventional estimates in the full 2005GV (option 1) and 2006GV (option 13); second (column 4) when the posterior coefficient of variation is larger than 1%. The latter case is based on the fact that posterior precisions in terms of coefficients of variation are typically much smaller than 1%.

The types of measured impacts are listed in column 5. They range from testing the impact of compilation assumptions, lesser data availability through surveys and administrative data sources, to lesser data availability on an annual basis. The impacts are measured by comparing the Bayesian estimates of alternative options, listed in column 6. For example, the impact of having less data annually for the current year 2006 is measured by comparing Bayesian estimates of options 14 and 13. Also, the impact of compiling only the SUT or only the IEA, can be assessed by comparing options 9 and 10 with option 1. A quantitative comparison between the alternative Bayesian results of these options is presented in the tables and sections referred to in column 7.

6.2 Tests of framework assumptions

The BSNA framework is very complex, because of the large number of variables, ratios, and identities. The application of the Bayesian estimation approach to such a complex framework was therefore carefully tested. The results of the most important tests are presented in Table 4.

TABLE 4

The first test was to check the internal consistency of ratios and identities in the framework. The test was carried out by applying SNAER and INTERFACE to the 2005GV benchmark data (option 1), and determine whether the Bayesian estimates are close to the conventional estimates. This should be the case, because data, data structures, and identities are fully compatible for that year. The only possible remaining distortion is the effect of the prior reliabilities. This is the additional information that was not (explicitly) available in the 2005 basic data compiled by conventional national accounts methods. The data in the first columns of Table 4 show that there are no distortions in the Bayesian estimates, and that in only three cases there are distortions for the posterior coefficients of variation. Two of these (household income taxes, paid and household other current transfers, received less paid; not shown in the table) are caused by the prior reliabilities, which may indeed be incompatible with the implicit reliabilities of the national accountants who did the benchmark compilation. A similar test for the full 2006GV framework, presented as option 13 in Table 6, shows slightly different results. In particular, the Bayesian estimates of gross fixed capital formation and changes in inventories (not shown in table) present significant changes from the conventional estimates. Also, changes in inventories have a posterior variation coefficient which is larger than 1%. The reason for these distortions may be that not all cells in the 2006GV scheme are treated as basic data. In particular, all output and import subitems of final and intermediate uses are missing (not estimated in conventional national accounting) and 2005 structures are used to estimate these variables.

The second test was to measure the impact of alternative values for the variables in the framework for which no basic data are available. The test was applied to the 2006YF framework. As explained earlier, values in cells without basic data are not used in the Bayesian estimation, except in the linearization of ratios. The results are shown in columns 3–6 of the table. In option 11, the 2005 benchmark data are included in the cells without basic data, and in option 14 so-called 'tentative' estimates are included in those cells. The latter are derived with the help of assumptions based on a selection of identities and ratios of 2005 structures; see also Section 6.3. The results of the test are convincing: In option 11 the number of distortions of estimates was 46, while in option 14 there were only 13 distortions. Also the number of distortions in coefficients of variation was less: a reduction from 6 to 3.

The third test was to determine whether cells without basic data (and therefore with formulas, YF) should be treated as variables without basic data, or alternatively as basic data, with reliabilities attached to those. The test was applied to the 2006 framework in options 14 (2006YF) and 15 (2006GF), and the results are presented in columns 7–10. It is clear that the alternative 2006YF is preferred, as it has less distortions than the 2006GF version (13 compared to 22) in the values of the estimates, while posterior coefficients of variation do not differ significantly between the two options. Based on tests 2 and 3 we conclude that option 14 (2006YF), in which formulas (F) are used to estimate the missing data (Y), is the preferred option for producing estimates for a current year.

The fourth test was to identify separately the impact of different Bayesian inputs, i.e. identities, price indices, and other ratios. The results are presented in options 7 and 8 (columns 11–14), which should be compared with option 1 in columns 1 and 2. When using only identities, the Bayesian estimates of all aggregates are, as expected, close to the conventional estimates (0.0% difference), while there are hardly any distortions in coefficients of variation. When adding price indices to this as an additional input, a few components of GDP by activity (manufacturing, construction, wholesale and retail trade) and household disposable income (mixed income gross and non-household operating surplus, gross, not shown in table) present significant deviations (more than 2%), while posterior coefficients of variation do not change significantly. When adding other ratios in option 1 the $> 2\%$ deviations in values of components of GDP and household disposable income disappear, but three distortions ($> 1\%$) in posterior coefficients of variation occur. Thus, price indices have large impacts, particularly on the coefficients of variation, but other ratios neutralize this impact. The significant price impact was confirmed in Table 8, where in option 5 of the 2006YF framework the reliability of price indices was significantly reduced from S (superior) to P (poor). When comparing the results of that option with those of option 14 (columns 10 and 11), the number of distortions of coefficients of variation increased significantly (from 3 to 11), but at the same time, somewhat unexpectedly, the number of distorted values of variables decreased (from 13 to 10).

The impacts of alternative options on five analytical indicators are also measured in this and subsequent tables: growth of GDP, implicit price deflator of total value added, propensity to consume of households (ratio of household final consumption and household disposable income), ratio of household disposable income and GDP, and the terms of trade change measured by the ratio between export and import price indices; see Reinsdorf (2010) for more comprehensive measures of terms of trade. Significant impacts were identified in option 11 (using 2005 values for variables not available in 2006) for the last three indicators, but not on GDP growth, and in option 15 (treating all values in 2006 as basic data) for the ratio of household disposable income to GDP.

6.3 Bayesian versus tentative estimates

In the previous subsection we saw that 2006 estimates improved if assumptions (selected ratios and identities) were used to estimate the variables for which no basic data were available. The estimates for variables without basic data, together with the basic data are called 'tentative' estimates. Using assumptions to arrive at values for the variables without basic data generates values that are more realistic than when assigning benchmark values to those cells, and this method is also close to the method used by national accountants.

FIGURE 2

The example presented in Figure 2, which represents a simple economy, may clarify the relation between tentative and Bayesian estimates. The figure includes three different versions of a data framework: The left panel is the framework for benchmark year t; the middle panel includes tentative estimates for the current year $t+1$; and the right panel is a framework with Bayesian estimates for the same current year $t+1$.

In each of the three frameworks there are 10 aggregates for output (P), imports (M), intermediate consumption (I), exports (X), value added (Y), gross fixed capital formation (K), final consumption (C), disposable income (R), domestic saving (S), and external saving (B). In addition, there are 6 identities and 5 ratios. In the left panel framework, the included identities are:

- Value added identity (E5 − E9 − E13 = 0),
- Supply-use identity ((E5 + K5) − (E9 + G9 + I9 + K9) = 0),
- Income distribution identity (E13 − I13 = 0),
- Saving identity (I13 − I9 − I16 = 0),
- Finance of capital formation identity (G9 − (I16 + K16) = 0),
- External deficit identity (K5 − K9 − K16 = 0),

and the included ratios are:

- input-output ratio (E9/E5),
- import-output ratio (K5/E5),
- capital-output ratio (G9/E5),
- propensity to consume ratio (I9/I13),
- export-output ratio (K9/E5).

All identities are equal to zero in the benchmark scheme, which means that the values of the 10 variables satisfy those identities. All ratio values in the benchmark scheme are considered to be the structural ratio values that do not only hold in the benchmark year, but can also be used as assumptions in the estimation for the current year.

The second scheme (middle panel) includes tentative estimates for all variables, based on available basic data and a selection of ratios and identities. In the derivation of tentative estimates, only those ratios that are used have the same value as in the base year, and only those identities that are used have a zero value in the second scheme. The basic values that are included in the second scheme are those for output (P), imports (M), exports (X), and gross fixed capital formation (K). Values for variables without basic data are derived with the help of assumptions, represented by a selection of ratios and identities, as follows: Intermediate consumption (I) is derived from output (P) with the help of the input-output ratio of the base year (E10); value added is derived from output (P) and intermediate consumption (I) with the help of the value added identity (P19); disposable income (R) is derived from value added (Y) with the help of the income distribution identity (U19); final consumption (C) is derived from disposable income (R) with the help of the propensity to consume ratio of the base year (I10); domestic saving is derived from disposable income (R) and final consumption (C) with the help of the saving identity (W19); and external saving (B) is derived as the difference between imports (M) and exports (X) with the help of the finance of external deficit identity (AB19).

Thus, four basic data, two ratios, and four identities (that is, precisely ten items of prior information) are used to arrive at tentative estimates of the ten variables in the current year $t + 1$. The estimates of the variables with basic data are equal to the values of those basic data, and the values of the ratios used for the tentative estimates are equal to the values in the first scheme. However, identities that are not used are not necessarily equal to zero (S19 supply-use and Y19 finance of capital formation identities), and ratios that are not used have values in the tentative estimates that are different from those of the base year (capital-output ratio U10, import-output ratio Y6, and export-output ratio Y10). Hence, the tentative estimates are not compatible with all identities and ratios of the scheme.

In the third scheme this incompatibility between estimates and the not-used identities and ratios has been repaired with the help of the Bayesian approach, which uses *all* ratios and identities. In addition, reliabilities of basic data and ratios are taken into account, so that basic data and ratio values that are less reliable are adjusted more than more reliable data and values. As a consequence, estimates differ between the second and third schemes. For example, we obtain I = 44, C = 34, K = 55, and X = 75 in the second scheme, and I = 41, C = 32, K = 56, and X = 75 in the third scheme. The differences between the values of variables between the second and third schemes are not large. Thus, small changes in the values of variables (and ratios), as compared to the second framework, make it possible to satisfy all identities, also those that were not satisfied in the second scheme.

The same principles are used in the main BSNA, when making estimates for a current year (2006): first tentative estimates and then Bayesian estimates. The last column of Table 3 shows which of the identities and ratios of the main BSNA scheme used in producing tentative estimates. They include most (within columns or vertical) behaviorist ratios (e.g. input-output coefficients, user coefficients, coefficients of components of household disposable income) in SUT and IEA, and exclude (across columns or horizontal) distributional ratios of value added in SUT and coefficients of distribution across revenue and expenditure items in IEA. We use vertically-defined identities, such as supply use identities, or identities defining value added and operating surplus in SUT and IEA, while horizontally-defined identities in IEA between revenues and expenditures are not used. This implies that in a scheme for the current year (2006) many of the identities that were not selected for use in the tentative estimates do not hold in the tentative estimates, and can only be satisfied when applying the Bayesian integration approach.

TABLE 5

The differences in the BSNA scheme of the tentative and Bayesian estimates from the conventional estimates of the full 2006GV scheme are presented in Table 5. There are slightly more distortions in the Bayesian than in the tentative estimates (13 versus 11), but the distortions concern different aggregates and are also smaller. Thus, the estimate of household disposable income is distorted in the tentative estimates and not anymore in the Bayesian estimates; the difference in the tentative estimates changes from 3.1% to 1.1% in the Bayesian estimates. Also the differences from conventional estimates between the Bayesian estimates of GDP in constant prices reduces considerably between tentative and Bayesian estimates; that of GDP in current prices only improved for the GDP total by expenditures. The same applies to the contribution to GDP in current prices of land transport and other taxes less subsidies on production; in both cases the differences from the conventional estimates are considerably reduced. Also the distortion in the estimate of the ratio of household disposable income to GDP reduces to a non-distorted value. Furthermore, the differences from conventional estimates reduce for the GDP growth rate and implicit price deflator of GDP. The deviation for the propensity to consume, however, increases. New distortions are also found in the Bayesian estimates of household final consumption and gross fixed capital formation. It is expected that if tentative estimates are improved by the national accountant, the differences from conventional estimates will become smaller not only for the tentative estimates but also for the Bayesian estimates.

6.4 Complete versus partial data in base year

The most important question is how well the Bayesian estimation method generates estimates that are close to conventional estimates, when annually only a partial data set is available. This question is examined on the basis of alternative sets of estimates presented in Table 6.

TABLE 6

The table compares Bayesian estimates for 2005 and 2006, alternatively based on full options 1 and 13, and limited annual data sets for those years (options

12 and 14). The limited data set corresponds to what is normally available annually. From this table several conclusions can be drawn.

When treating a large number of cells in the 2005 framework as variables for which annually no data are available (2005YF, option 12), Bayesian estimates of some GDP expenditures (gross fixed capital formation and also exports and imports) and some components of household disposable income (not shown in the table) differ significantly ($> 2\%$) from the conventional estimates. Bayesian estimates of GDP totals, however, differ only slightly more for the 2005YF framework than for the 2005GV framework, but the total of household disposable income deviates significantly from its conventional estimate in the 2005YF framework. Posterior coefficients of variation do not differ significantly between the 2005GV and 2005YF options, and in both cases are much smaller than their prior equivalents.

A similar pattern is observed when comparing 2006YF and 2006GV estimates (options 14 and 13), but some differences should be noted. Bayesian estimates of GDP are slightly better for 2006YF than for the 2006GV version. The opposite is true for household disposable income, but the deviation from the conventional estimates is significantly less in the 2006YF framework than in the 2005YF framework. There are also differences between the 2005YF and 2006YF frameworks with regard to the components of GDP and household disposable income.

When comparing the estimates of the 2005YF and 2006YF (options 12 and 14), we see that in the 2005YF framework there are no significant deviations in the industry breakdown of GDP, while in the 2006YF framework contributions to GDP of construction, wholesale and retail trade, and land transport show significant deviations from the conventional 2006 estimates. For the expenditures the distortions in the 2005YF framework are in gross fixed capital formation, exports, and imports; in the 2006YF framework they are in household and non-profit institutions final consumption and in gross fixed capital formation. With regard to the IEA variables explaining household disposable income, the deviations from conventional estimates are almost the same in the 2005YF and 2006YF frameworks. There are no significant differences in the coefficients of variation between the 2005YF and 2006YF frameworks. Thus, aggregates of GDP in current and constant prices, when estimated with limited annual data, show insignificant ($< 2\%$) differences from the 2005 and 2006 conventional estimates, while the aggregate of household disposable income differs more from the conventional estimates in the 2005YF and 2006YF options, and is therefore more dependent than GDP on the availability of basic data. Components of GDP by activities and expenditures, and also components of household disposable income, are more dependent on the availability of basic data than the totals.

It should be noted that there is hardly any effect of limited data availability on the measurement of GDP growth in 2006, the implicit price deflator of total value added, the propensity to consume of households, and the terms of trade effect. As the effect on household disposable income in the 2006YF framework is much smaller than in the 2005YF framework, the impact on the GDP/household disposable income ratio is also much lower in the 2006YF framework than in the 2005YF framework.

7 Sensitivity analysis

7.1 Sensitivity to SUT or IEA scope

Many countries limit their SNA compilation to the SUT. At the other extreme, we may consider compilation of only the IEA, which may be relevant if analysts are interested in sector accounts and their monetary extension. The question raised here is how reliable estimates are if they are based solely on the compilation of the SUT or alternatively of the IEA.

TABLE 7

The Bayesian estimates in Table 7 answer this question for the 2005GV framework of the benchmark year, by comparing options 9 and 10 to option 1. The results show that if only the SUT is compiled (option 9), the number of distortions in the estimates remains 0, as in option 1, while the number of distortions of the coefficients of variation increases slightly (from 3 to 6). Hence the precision of the compilation of the SUT is nearly the same as when the whole framework is compiled. This is not the case, however, when only the IEA is compiled. In that case, unexpectedly perhaps, Bayesian estimates of government final consumption deviate considerably from the conventional estimates, and the same holds (not shown in the table) for household social transfers, received (+) less paid (−). The number of distortions in the posterior coefficient of variation increases dramatically from 3 to 15, if only the IEA is compiled. The latter is to be expected, because much fewer basic data and also identity and ratio restrictions are used in this compilation.

7.2 Sensitivity to data available from surveys and administrative data sources

The next question is whether some basic data influence the precision of the estimates more than others, i.e. which surveys and administrative data should be considered essential for compiling reliable national accounts aggregates. In Table 7 we quantify the effect for the 2006YF framework: in option 16 for data availability on household final consumption data (columns 10 and 11), in option 17 for data on exports and imports (columns 12 and 13), in option 18 for data on services (columns 14 and 15), and in option 19 for data on financial corporations (columns 16 and 17). In all cases, impacts are assessed by comparing the distortions in the Bayesian estimates and the posterior coefficients of variation with those for option 20 in columns 8 and 9. Option 20 is close to option 14, but includes all final consumption restrictions.

The impact of household final consumption restrictions is measured by comparing option 16 (without restrictions) to option 20 (with restrictions). The absence of the restrictions in option 16 consists in not using the household survey data, the assignment of a poor (P) reliability to the weights of household final consumption items in the benchmark household survey, and also the assignment of a poor (P) reliability to the use coefficients (in order to eliminate the impact of structural coefficients that may mitigate the influence of not having household survey data and ratios). The overall impact is limited: Without using household survey data, the number of distortions in the values of the Bayesian estimates increases from 11 to 13, while the number of distortions in the coefficients of variation remains the same (2). The additional distortions are mainly in the two items of household final consumption in current and constant prices, as expected. There are some shifts in distortions for individual activity categories of GDP in current and constant prices, but overall there are no serious distortions in the main aggregates of GDP in current and constant prices and household disposable income, and in the details thereof.

Not having import and export details would be the case in a regional economy in which it is difficult to register incoming and outgoing flows of goods, in countries where foreign trade statistics are little developed, or in countries belonging to a customs union such as the European Union. This situation is simulated in option 17, which assigns poor (P) reliability to import and export data, balance of payments, and external sector data, and also, as before, to the user coefficients. To measure the impacts, a comparison is made between the distortions in options 17 and 20. The impacts in this case are also limited, and are mainly found in the four items of exports and imports in current and constant prices, leading to 15 distortions in the values of the estimates in option 17 against 11 in option 20. The main aggregates of GDP in current and constant prices and household disposable income are not in any major way affected by the lack of export and import and external sector data. The number of distorted values of coefficients of variation in this case is higher than in option 20: There are 9 instead of 2 distorted values of those coefficients, and the additional ones are in exports and imports in current and constant prices, and in several components of household disposable income (not shown in the table).

Impacts are much larger in a third scenario, in which it is assumed that there are no reliable statistics on services. This is simulated in option 18, in which data on services are assigned poor (P) reliability. When comparing the distortions in option 18 with those in option 2, we observe that the number of distortions in the values of the Bayesian estimates increases from 11 to 32, while the number of distortions in the coefficients of variation remains the same (2). The distortions in the values of the Bayesian estimates occur in all three major aggregates of GDP in current and constant prices and also in household disposable income. Several subcomponents of the three aggregates are affected. Posterior coefficients of variation are not affected.

The impact of not having reliable data on financial corporations was simulated in option 19. In this option it was assumed that all data on financial corporations in the SUT and IEA had poor (P) reliability. When comparing the distortions in this option 19 with those in option 20, we found that the number of distortions in the Bayesian estimates increased significantly from 11 to 23, while the number of distortions in the posterior coefficients of variation increased slightly from 2 to 4. In this case Bayesian estimates of the main aggregates of GDP in current and constant prices and household disposable income are not significantly distorted. The distortions only occur in the subcomponents of GDP by activities that are related to services, and also in several subcomponents of household disposable income (not shown in the table). The additional distortions in the coefficients of variation also occur in these subcomponents.

7.3 Sensitivity to aggregation

The present framework with 2719 variables is a very large one. The number of aggregate variables is, however, relatively small (62), and a subset of these are presented in the tables. This follows national accounts practices, which is generally also carried out in much detail leading to a small set of estimates, sometimes just one: GDP. This procedure is based on the assumption that it is more precise to estimate aggregates by using much detail in the compilation. Added complexity does not *necessarily* improve the estimates, and this is tested in option 21 in the last column of Table 7.

To carry out this test a small framework was designed, containing the main aggregates of GDP in current and constant prices and household disposable income, but also including subcomponents that are needed to link these variables conceptually and quantitatively. The main aggregates are output, intermediate consumption, value added, value added components classified by aggregate ISIC categories, and total imports and exports, as well as sector components that link GDP to household disposable income. The total number of variables in the aggregate framework is 105, of which 29 (27.6% versus 19.5% in the extended framework) are supported by basic data, and there are 29 identities and 83 ratios. The number of information items (basic data, identities, and ratios) per variable is therefore 1.34 (versus 1.45 in the extended BSNA framework). Reliabilities are attached to basic data and ratio values. Although similar in nature, these reliabilities are not quantitatively linked to those of the extended framework presented in Tables 1 and 2.

Column 18 in Table 7 shows that the number of distortions in the estimated values increases to 34, when an aggregate 2006YF framework is used, as compared to 11 distortions in option 21 for the extended 2006YF framework. Distortions in the coefficients of variation are not presented for the aggregate framework, as they are not comparable with those of the extended framework. We conclude that Bayesian estimates compiled in a disaggregated framework are generally closer to the conventional estimates than those compiled through aggregates.

There is another aspect of aggregation, the results of which are reflected in Tables 4–8. In several options in those tables, it is shown that the values and coefficients of variation of GDP and household disposable income are generally not distorted, while distortions are observed for components of GDP and household disposable income. For example, in option 14 of the 2006YF framework in Table 6, there are no significant distortions in the values of GDP in current and constant prices and household disposable income, and also the growth rate is not much affected. However, for some details of these aggregates by industries and other categories, there are significant distortions, such as for contributions to GDP of electricity, gas and water and construction, household final consumption and gross fixed capital formation.

7.4 Sensitivity to prior precisions

Prior reliabilities play an important role in the Bayesian estimation procedure. The prior coefficients of variation are subjectively determined by national accountants based on insight and some detailed studies; see also Bos (2009). We now ask how much influence these priors have on the Bayesian estimates and how much posterior coefficients of variation are reduced when compared to their prior values. We will also consider the question to what extent values of Bayesian estimates of variables with basic data are located outside the range of the basic values plus/minus a percent variation that is permitted by the national accountant, when quantifying the prior coefficient of variation.

TABLE 8

In Table 8 a comparison is made between posterior coefficients of variation of the full 2005GV benchmark using standard prior reliabilities and the same framework in which these standard prior reliabilities are changed. The same exercise is carried out for the 2006YF framework with limited data availability. For the 2005GV framework the impact is measured by increasing or decreasing the standard prior reliabilities proportionally with ±50% and by using fictitious prior reliabilities of ±100% for all basic data and ratio values. For the 2006YF framework the impact is measured by lowering the S (superior) reliability of price indices to P (poor) and by increasing proportionally all coefficients of variation by 50%.

The impact of the changes in prior reliabilities on the estimates is low. In the 2005 framework (options 2, 3, 4 versus 1) the number of distortions in estimates (i.e. $> 2\%$ change as compared to conventional estimates) remains zero, even when fictitious prior reliabilities are used. A similar pattern is observed for the 2006YF framework (options 5 and 6 versus 14). In that framework the number of distortions in estimates of variables remains the same (13), and when the reliability of price indices is lowered from S to P, the number of distortions in estimates is even reduced to 10.

The change in the prior reliabilities does, however, have a significant impact on the posterior coefficient of variation of the estimates. In the 2005GV framework, this effect is largest in the case of using fictitious prior coefficients of variation (64 instead of 3 distortions), a little smaller when proportionally increasing the coefficients of variation (11 distortions), and hardly any impact when proportionally decreasing the prior coefficients of variation (4 distortions).

In the case of the 2006YF framework, the impact on posterior coefficients of variation is insignificant, when prior coefficients of variation are increased proportionally with 50%. Lowering the reliability of price indices in the 2006YF framework has no (or even a decreasing) effect on the number of distortions in the Bayesian estimates, but the number of distortions in the posterior coefficients of variation has increased (from 3 to 11).

In Table 8, as in earlier tables, we note the significant change in the posterior reliability as compared to the prior reliability of variables. In most instances, prior reliabilities of 3, 5, and 12% change to much lower posterior reliabilities of less than 1%, independent of their prior value. In the 2006YF framework of option 14 the variation coefficient of GDP in current prices changes from 3% prior to 0.01% posterior, and in constant prices from 3% to 0.07%. For household disposable income the reduction is from 12% to 0.08%. Posterior coefficients of variation are generally higher for details than for the totals of GDP and household disposable income, and also higher when using only the IEA in the estimation of the two aggregates (option 10 in Table 7). We emphasize that, while posterior variances are always smaller than prior variances (in accordance with Bayesian theory), this is not necessarily the case for coefficients of variation. In fact, we found that in 38% of the cases (1173 of the 3053 variables estimated), the deviation of posterior estimates from conventional estimates deviated more than the percent value of the prior coefficient of variation. Overall, the findings confirm the utility of using a Bayesian approach to national accounts integration; the approach makes the posterior estimates much more reliable.

8 Conclusions

Based on experiences and sensitivity experiments with the Bayesian estimation approach as described in this paper we draw eight conclusions.

After many INTERFACE/SNAER runs applied to various versions of the 2005 and 2006 framework of Guatemala and other countries in Central America, the BSNA framework and the Bayesian estimation approach can be considered as operational, not only in Guatemala but, with minor modifications, in all countries using the SNA. In other words, the software programmes can be applied to any framework of data and variables, for which identity and ratio relations and reliabilities can be defined and presented in EXCEL format, leading to a consistent set of Bayesian estimates and posterior coefficients of variation. The selection of basic data, identities, ratios, and reliabilities are essential ingredients in the Bayesian approach presented here. As long as these four elements can be well identified, the Bayesian estimation method does not compete but can be well combined with alternative methodologies such as ERETES developed by EUROSTAT, and the method proposed in Rueda-Cantuche and Ten Raa (2009) to construct input-output tables or the entropy method used in Robilliard and Robinson (2003) to reconcile household survey data and national accounts.

It proved feasible to define an internally consistent and very large BSNA framework of 2719 variables, including not only the SUT and the major aggregate of GDP, but also the IEA which incorporates household disposable income as another major aggregate. The internal consistency of the framework means that there is full compatibility of Bayesian inputs, and this is confirmed by applying INTERFACE and SNAER to the 2005GV framework, resulting in Bayesian estimates that are close to the conventional estimates and without any distortions in the posterior coefficients of variation. The internal consistency of the framework is also reflected in estimates of price indices that remain close to 1.00 in the base year.

The use of assumptions (i.e. a selection of ratios and identities) to arrive at tentative estimates for all variables of the framework, is found to be the best procedure to prepare the framework for Bayesian integration and arrive at current year estimates (2006YF framework). It leads to Bayesian estimates for a current year that are close to conventional estimates for that year, in particular for major aggregates such as GDP and household disposable income. More efforts are needed by national accountants to improve the tentative estimates prior to Bayesian integration.

A set of basic data for a current year (2006YF) on output, imports and exports, employment, data on the government and financial corporate sectors, and the rest of the world, constituting approximately 20% of the total number of 2719 variables in the system, is adequate for making Bayesian estimates of the comprehensive framework of SUT and IEA, as long as the basic data are supplemented by a large number of identities and ratio values including a full set of price indices. This would apply not only to the compilation of annual accounts, but may also hold for early estimates, quarterly accounts, or even projections. When the number of basic data gradually increases between early (flash), preliminary, semi-final, and final estimates, while the number of identities and ratios remains the same, Bayesian estimates gradually improve, i.e. discrepancies between Bayesian and conventional estimates reduce and the number of estimates with distorted posterior coefficients of variation also reduces. In the present framework for annual accounts the ratio of the number of information items (basic data, ratios and identities) and the number of variables is 1.45, while the maximum ratio (in this framework) is 2.02 when basic data are available on all variables.

Compilation through only the SUT yields Bayesian estimates that are close to those of the full framework. This means that the IEA data and their restrictions do not add much to the information that is needed to arrive at GDP and other main aggregates. Compilation through the IEA only is much less reliable.

The best procedure to compile major aggregates is through the compilation of details. Compilation on the basis of aggregate basic data is found to be much less reliable. Within this preferred procedure major aggregates are compiled much more reliably than details by economic activities and sectors and by expenditure and income components of GDP and household disposable income.

By analyzing the sensitivity of Bayesian estimates to the availability of basic data from selected survey and administrative data sources, we find that direct data on the financial corporate sector and on services have most impact. Direct information from household surveys on household final consumption and from foreign trade and balance of payments statistics on exports, imports, and external sector transactions has less impact.

The posterior coefficients of variation are much smaller than the prior values. This means that integration of estimates, as being pursued by the Bayesian integration method in line with similar practices by national accountants, results in a considerable improvement of the reliability of Bayesian (and also conventional) estimates. The reduction in posterior coefficients of variation is the most in household disposable income and GDP in current prices, less in GDP in constant prices, and even less in details of both aggregates.

In addition to these eight conclusions, we mention a few thoughts for future work. Regarding the use of basic data, more work is required. We need data that are 'basic', hence not already processed. In practice, this is not fully possible, because all data are processed. But there are degrees of processing, and the more basic the better in our method; see also the distinction made in Vanoli (2010) between the use of national accounts frameworks for observation and analysis.

Regarding the Bayesian method, more calibration of Bayesian estimates with conventional estimates is needed in order to specify the ideal features (ratios, identities, and reliabilities) of a framework of this size, in which Bayesian estimates are close to conventional estimates, not only for aggregates but also for details. More experiments are also needed in order to arrive at estimates for current years (later than 2006) that are further away from the base year (2005). The normal distribution has its limitations, especially the symmetry implied by this distribution. The possibility of introducing inequalities in the estimation should be investigated; see Boonstra *et al.* (2010). Finally, the possibility of using parameters of multivariate regression functions instead of the present binary ratios would be of interest.

Regarding the framework, alternatives may be considered. The present use is in annual accounts, but work is already underway to apply our methods to quarterly accounts. The use of frameworks for alternative monetary analysis has been applied to regional accounts, their use in monetary analysis in which the IEA is extended to financial accounts is being investigated, and also their use in projections. Furthermore, instead of only using monetary variables as in the present BSNA framework, non-monetary variables may be introduced in satellite frameworks. Efforts are underway in defining and implementing a Bayesian integration of health-environmental satellite frameworks with (non)monetary variables, and frameworks for demographic analysis, as in Gross *et al.* (2009). In general, any study based on large data sets may be supported by specialized data frameworks, in which the data of the study are made compatible with each other through Bayesian integration, and missing variables are estimated as well.

References

Boonstra, H.J., C. de Blois, and G.J. Linders (2010). Macro-integration with inequality constraints: An application to the integration of transport and trade statistics, Discussion Paper 10006, Statistics Netherlands, The Hague/Heerlen.

Bos, F. (2009). The art and craft of compiling national accounts statistics and their implications for reliability, *The Review of Income and Wealth*, 55, 930–958.

Danilov, D. and J.R. Magnus (2007). Some equivalences in linear estimation (in Russian), *Quantile*, 3, 83–90. (English original can be downloaded from http://cdata4.uvt.nl/websitefiles/magnus/paper77b.pdf.)

Danilov, D. and J.R. Magnus (2008). On the estimation of a large sparse Bayesian system: The Snaer program, *Computational Statistics & Data Analysis*, 52, 4203–4224.

Gould, N.I.M. and J. Nocedal (1998). The modified absolute-value factorization for trust-region minimization. In: R. De Leone, A. Murli, P.M. Pardalos, and G. Toraldo (eds.), *High Performance Algorithms and Software in Nonlinear Optimization*, Kluwer Academic Publishers, Dordrecht, 225–241.

Gross, K., B.A. Craig, and W.D. Hutchison (2009). Bayesian estimation of a demographic matrix model from stage frequency data, *Ecology*, 83, 3285–3298.

Jorgenson, D.W. (2009). A new architecture for the U.S. national accounts, *The Review of Income and Wealth*, 55, 1–42.

Jorgenson, D.W., S. Landefeld, and W. Nordhaus (2010). A new architecture for the U.S. national accounts: A reply to André Vanoli, *The Review of Income and Wealth*, 56, 752–761.

Magnus, J.R. and J.W. Van Tongeren (2002). Assessing and optimizing the impact of national accounts compilation methods on indicator analysis, in: *United Nations Statistics Division: Handbook of National Accounting: Use of Macro Accounts in Policy Analysis*, Studies in Methods Series F, No. 81, Chapter VI-D, 263–281, United Nations: New York.

Magnus, J.R., J.W. Van Tongeren, and A.F. De Vos (2000). National accounts estimation using indicator ratios, *The Review of Income and Wealth*, 46, 329–350.

Rao, C.R. (1971). Unified theory of linear estimation, *Sankhyā, A*, 33, 371–477. (Corrigenda, *Sankhyā, A*, 34, 477.)

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition, John Wiley, New York.

Reinsdorf, M.B. (2010). Terms of trade effects: Theory and measurement, *The Review of Income and Wealth*, 56, S177–S205.

Robilliard, A.-S. and S. Robinson (2003). Reconciling household surveys and national accounts data using a cross entropy estimation method, *The Review of Income and Wealth*, 49, 395–406.

Rueda-Cantuche, J.M. and T. ten Raa (2009). Testing the assumptions made in the construction of input-output tables, International Input-Output Association, Working Paper WPIOX 09-011.

Stiglitz, J.E., A. Sen, and J.-P. Fitoussie (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress, www.stiglitz-sen-fitoussi.fr.

United Nations, European Commission, OECD, IMF, and World Bank (1993). *System of National Accounts 1993*, Sales No. E.94.XVII.4, United Nations: New York.

United Nations, European Commission, OECD, IMF, and World Bank (2008). *System of National Accounts 2008*, Sales No. 08.XVII.29, United Nations: New York.

United Nations Statistics Division (UNSD) (1989). *Classification by Broad Economic Categories, Defined in terms of SITC, Rev. 3 (BEC Rev. 3)*, Series M, No. 53, United Nations: New York.

Van Tongeren, J.W. (2004). Designing analytical data frameworks, *The Review of Income and Wealth*, 50, 279–297.

Van Tongeren, J.W. (2010). *Bayesian Integration of Large SNA Data Frameworks*, Final report of BSNA Application in CMCA Region, Delft, The Netherlands.

Vanoli, A. (2010). The new architecture of the U.S. national accounts and its relationship to the SNA, *The Review of Income and Wealth*, 56, 734–751.

Table 1: Basic data in the BSNA framework

	Origin of data	Prior reliability
Output data	Output data in early estimates generally is available in constant prices, as derived from growth rates based on data from surveys that request data on output in physical units of selected establishments. The constant price data are then converted to current price data, following established procedures of national accountants. Both current and constant price data on output are included in the BSNA framework, even though only one of the data sets in current or constant prices is strictly needed, as the other can be easily derived with help of price indices.	**F** for total and detail by products
Import and export data	Import and export data are generally available in detail from foreign trade statistics (goods) and Balance of Payments (services).	**F** for total, **H** for detail by products
HH final consumption data	HH final consumption data are generally only available directly in years in which HH surveys are conducted. However, national accounts practices update the results of those surveys to future periods, using very detailed information on products, consumption weights (of CPI) and changes in the population between the years to arrive at revised estimates. As this detailed updating of the HH final consumption data cannot be well reconstructed in the present Bayesian estimation procedure (ratios and identities), the result of the national accounts updating is directly included in the BSNA framework as basic data.	**H** for total, **M** for detail by products
Employment data	Employment data are generally available from employment surveys, either or not included in HH surveys. National accountants often update this information to recent periods, using extrapolation based on demographic statistics.	**F** for total and **M** for detail by economic activities
GOV sector data	GOV sector data are directly based on data available in the GOV administrative records.	**F**
FC sector data	FC data are generally available from Central Bank reports (administrative data source) about the Central Bank and other banks and insurance companies. Other finanancial institutions, such as currency exchange houses, insurance agents and stock brokers, are generally estimated additionally by national accountants to complete the FC sector. All data including the additional estimates made by the national accountants are treated as basic data in the BSNA framework .	**F**
External sector data	External sector data are covered through periodic updates of the Balance of Payments by the Central Bank. Their conversion to SNA format is used to complete the basic data cells of this sector in the BSNA.	**F**
Price indices	Price indices are partly based on surveys of consumer prices, retail and wholesale trade surveys, producers' price surveys. The remaining part of the indices is constructed by national accountants on the basis of these survey data. Both the price data from surveys and also the estimates made by national accountants are included in the BSNA framework as very reliable ratio information.	**F**

Table 2: Identities and ratios (other than price indices) in the BSNA framework

	Identities and ratios described	Prior reliability	Used in tentative estimates
SUT-identities	Supply-use identities in current and constant prices for output and imports separately		X
	Identities between output by industries and product and secondary production		
	Trade and transport identity, also in current and constant prices, between margin and output of trade and transport		
	Identities of value added and operating surplus in current and constant prices		X
	Identities of HH final consumption by product between output and imports allocated to HH final consumption, and HH survey data in current prices		
	Identities between totals of uses and import and output components		X
	Identity between intermediate consumption by industry of use and product in current and constant prices		
	Identity between gross fixed capital formation by industry of use and product in current and constant prices		
	Identities expressing totals as sum of details by CPC and ISIC categories in current constant prices		X
	GDP identities		
IEA-identities	Identities between receipts and expenditures of sectors		
	Identities defining in IEA: value added, operating surplus, GOV & NPI final consumption, social transfers, disposable income, saving, net lending		X
SUT-IEA-identities	Identities between industry and sector production accounts data		
SUT-IEA-ratios	Coefficients of distribution of industry production account data to sectors	**S**	X
SUT behaviorist or vertical ratios	Input output ratios in constant prices between (i) intermediate consumption and output, (ii) gross fixed capital formation by industries of use and output and (iii) employment and output (labour-productivity coefficient)	**S**	X
	SUT secondary product ratios	**H**	X
	User coefficients of output and imports	**H**	X
	Wage and mixed income rates	**M**	
	Product tax rates	**H**	X
	Trade and transport margins	**H**	X
SUT distributional ratios	Coefficients of distribution of HH final consumption by products in SUT	**S**	
	Coefficients of distribution of value added and product taxes between industries and products	**S**	
IEA behaviorist or vertical ratios	Finance of investment ratios	**M**	X
	Coefficients of components of HH disposable income	**M**	X
	Propensity to consume of HH	**M**	X
	Income tax ratios by sector	**M**	X
IEA horizontal distributional ratios	IEA coefficients of distribution of revenues and expenditures between sectors	**S** for disposable income, saving and net lending **L** for other IEA distribution coefficients	
SUT-IEA aggregate ratios	GDP growth rate Propensity to consume HH disposable income/GDP ratio PX/PM (Terms of trade)		

Table 3: Impact of alternative compilation options on Bayesian estimates

Compilation options	Bayesian conditions of options	Nr. of distortions in Bayesian estimates		Impact of options on Bayesian estimates		
		Nr. of estimates >2% change	Nr. of estimates with posterior coefficients of variation >1%	Types of impacts measured	Impact measured by comparing following options	Reference to tables and sections, where impacts are quantified/ assessed
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Option 1	2005GV	0	3	Impact of changes, i.e. increases and decreases of prior coefficients of variation, with full 2005 benchmark data	comparing options 2, 3 and 4 with 1, and also reviewing options 5 and 6	Table 8, section 7.4
Option 2	2005GV, 50% decrease of prior coefficients of variation	0	4			
Option 3	2005GV, 50% increase of prior coefficients of variation	0	11			
Option 4	2005GV, fictitious prior coefficients of variation (=100%)	0	64			
Option 5	2006YF, poor (P) reliability of price indices	10	11			
Option 6	2006YF, 50% increase of prior coefficients of variation	13	5	Impact of increases of prior coefficients of variation, with limited 2006 annual data		
Option 7	2005GV, only identities included	0	1	Impact of Bayesian conditions, i.e. identities, price indices, other ratios	comparing options 7 and 8 with 1	Table 4, section 6.2
Option 8	2005GV, identities and price indices included	10	1			
Option 9	2005GV, SUT only	0	6	Impact of only compiling SUT or IEA	comparing options 9 and 10 with 1	Table 7, section 7.1
Option 10	2005GV, IEA only	2	15			
Option 11	2006YV(2005), 2005 values for missing data	46	6	Impact of using assumptions, i.e. ratios and identities to derive tentative values for missing data	comparing options 14 and 11	Table 4, section 6.2
Option 12	2005YF	15	3	Impact lesser availability of basic data in 2005 benchmark year	comparing option 12 with 1	Table 6, section 6.4
Option 13	2006GV	5	5	Impact lesser availability of basic data in 2006 current year	comparing options 13 and 14	
Option 14	2006YF	13	3			
Option 15	2006GF, tentative values for missing data treated as basic data	22	2	Impact of assigning basic data status to tentative estimates	comparing options 14 and 15	Table 4 section 6.2
Option 16	2006YF, poor quality of household final consumption data and ratios	13	2	Impact of poor quality/non-availability respectively of data on HH final consumption, imports-exports and external sector, services, financial corporations	comparing options 16, 17, 18, 19 with option 20	Table 7, section 7.2
Option 17	2006YF, poor quality of imports, exports and external sector data	15	9			
Option 18	2006YF, poor quality data on services	32	2			
Option 19	2006YF, poor quality of data on financial corporations	23	4			
Option 20	2006YF, option 14 with full household final consumption restrictions	11	2			
Option 21	2006YF, aggregate variables and basic data, adjusted aggregate reliabilities	34	Not comparable	Impact of using aggregates in compilation	comparing options 21 with 14	Table 7, section 7.3

Table 4: Testing alternative framework assumptions

	Conventional estimates 2005	Conventional estimates 2006	Testing the application to benchmark data (2005GV)		Testing alternative values for variables without basic data (2006YF)				Testing alternative scopes of basic data (2006)				Testing the accumulative use of Bayesian inputs (2005GV)			
			Bayesian estimates of 2005GV option 1		Using values of base year option 11		Using assumptions (selected ratios and identities) to derive tentative values option 14		Treating only selected variables as variables with basic data (2006YF) option 14		Treating all variables as variables with basic data (2006GF) option 15		Identities only option 7		Identities and price indices option 8	
			% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Nr. of distorted values			*0*	*3*	*46*	*6*	*13*	*3*	*13*	*3*	*22*	*2*	*0*	*1*	*10*	*1*
GDP ECONOMIC ACTIVITIES, CONSTANT			**0.0%**	**0.1%**	**1.2%**	**0.10%**	**0.0%**	**0.07%**	**0.0%**	**0.07%**	**-1.4%**	**0.07%**	**0.0%**	**0.00%**	**1.5%**	**0.00%**
GDP EXPENDITURES, CONSTANT			**0.0%**	**0.1%**	**1.2%**	**0.10%**	**0.0%**	**0.07%**	**0.0%**	**0.07%**	**-1.4%**	**0.07%**	**0.0%**	**0.00%**	**1.5%**	**0.00%**
Agriculture, forestry and fishing			0.0%	0.1%	13.1%	0.23%	1.4%	0.10%	1.4%	0.10%	1.1%	0.10%	0.0%	0.01%	0.3%	0.01%
Mining and quarrying			0.0%	0.1%	21.5%	0.16%	-0.1%	0.12%	-0.1%	0.12%	-0.1%	0.12%	0.0%	0.05%	0.1%	0.06%
Manufacturing			0.0%	0.2%	-15.9%	1.04%	-0.6%	0.26%	-0.6%	0.26%	-3.8%	0.25%	0.0%	0.02%	3.3%	0.02%
Electricity, gas and water supply			0.0%	0.3%	35.9%	0.39%	2.4%	0.36%	2.4%	0.36%	1.2%	0.35%	0.0%	0.02%	1.2%	0.01%
Construction			0.0%	0.5%	20.6%	0.40%	-7.8%	0.57%	-7.8%	0.57%	-7.9%	0.55%	0.0%	0.01%	4.7%	0.00%
Wholesale and retail trade			0.0%	0.2%	4.5%	0.27%	1.9%	0.22%	1.9%	0.22%	-2.2%	0.22%	0.0%	0.00%	2.5%	0.00%
Land transport			0.0%	0.5%	-28.9%	0.88%	1.5%	0.47%	1.5%	0.47%	1.0%	0.46%	0.0%	0.01%	1.3%	0.01%
Transport (other), storage, telecommunication			0.0%	0.3%	38.1%	0.25%	-1.1%	0.29%	-1.1%	0.29%	-2.5%	0.29%	0.0%	0.02%	1.3%	0.02%
Financial and insurance activities			0.0%	0.0%	-8.0%	0.25%	-0.4%	0.11%	-0.4%	0.11%	-0.1%	0.03%	0.0%	0.02%	0.0%	0.02%
Private services			0.0%	0.0%	-2.3%	0.30%	-0.1%	0.02%	-0.1%	0.02%	-0.7%	0.02%	0.0%	0.01%	0.5%	0.01%
Public administration and defence; compulsory social security			0.0%	0.0%	2.3%	0.06%	-0.4%	0.04%	-0.4%	0.04%	0.2%	0.02%	0.0%	0.01%	0.0%	0.01%
VALUE ADDED GROSS, TOTAL, CURRENT			**0.0%**	**0.0%**	**2.3%**	**0.00%**	**-0.1%**	**0.00%**	**-0.1%**	**0.00%**	**-1.6%**	**0.00%**	**0.0%**	**0.00%**	**1.6%**	**0.00%**
GDP ECONOMIC ACTIVITIES, CURRENT			**0.0%**	**0.0%**	**2.2%**	**0.05%**	**-0.1%**	**0.01%**	**-0.1%**	**0.01%**	**-1.7%**	**0.00%**	**0.0%**	**0.00%**	**1.5%**	**0.00%**
Household final consumption			0.0%	0.0%	15.3%	0.09%	2.7%	0.00%	2.7%	0.00%	-0.2%	0.00%	0.0%	0.00%	2.0%	0.00%
Government final consumption			0.0%	0.0%	-0.2%	0.01%	-0.2%	0.01%	-0.2%	0.01%	0.2%	0.01%	0.0%	0.00%	0.0%	0.00%
NPISH final consumption			0.0%	0.4%	1.7%	0.47%	0.5%	0.45%	0.5%	0.45%	-0.3%	0.11%	0.0%	0.03%	-0.1%	0.02%
Gross fixed capital formation			0.0%	0.2%	0.0%	3.21%	-5.4%	0.20%	-5.4%	0.20%	-3.1%	0.18%	0.0%	0.00%	-1.3%	0.00%
Exports			0.0%	0.0%	2.9%	0.09%	-0.7%	0.09%	-0.7%	0.09%	-0.3%	0.05%	0.0%	0.00%	0.0%	0.00%
Minus: Imports			0.0%	0.0%	-5.3%	0.06%	1.1%	0.05%	1.1%	0.05%	0.3%	0.03%	0.0%	0.00%	0.0%	0.00%
GDP EXPENDITURES, CURRENT			**0.0%**	**0.0%**	**2.2%**	**0.05%**	**-0.1%**	**0.01%**	**-0.1%**	**0.01%**	**-1.7%**	**0.00%**	**0.0%**	**0.00%**	**1.5%**	**0.00%**
Compensation of employees			0.0%	0.1%	-38.0%	0.24%	0.7%	0.14%	0.7%	0.14%	-0.9%	0.13%	0.0%	0.00%	0.4%	0.00%
Other taxes less subsidies on production			0.0%	0.1%	-98.5%	0.65%	0.6%	0.10%	0.6%	0.10%	4.6%	0.07%	0.0%	0.02%	0.0%	0.02%
Non-HH operating surplus, gross			0.0%	0.2%	44.0%	0.11%	-0.6%	0.18%	-0.6%	0.18%	-1.0%	0.17%	0.0%	0.00%	2.8%	0.00%
HH Operating surplus (on dwelling services), gross			0.0%	0.1%	4.4%	0.13%	0.2%	0.14%	0.2%	0.14%	0.0%	0.14%	0.0%	0.00%	0.0%	0.00%
Mixed Income, gross			0.0%	0.2%	-1.4%	0.33%	-0.8%	0.22%	-0.8%	0.22%	-4.6%	0.22%	0.0%	0.00%	2.1%	0.00%
VALUE ADDED GROSS, TOTAL, CURRENT			**0.0%**	**0.0%**	**2.3%**	**0.00%**	**-0.1%**	**0.00%**	**-0.1%**	**0.00%**	**-1.6%**	**0.00%**	**0.0%**	**0.00%**	**1.6%**	**0.00%**
HOUSEHOLD DISPOSABLE INCOME, GROSS	**Conventional estimates**		**0.0%**	**0.1%**	**9.5%**	**0.09%**	**1.1%**	**0.08%**	**1.1%**	**0.08%**	**0.4%**	**0.07%**	**0.0%**	**0.00%**	**1.7%**	**0.00%**
	2005	**2006**														
GDP growth rate 2006/2005		**5.380%**			**1.3%**		**0.0%**		**0.0%**		**-1.5%**					
Implicit price index Value Added, total	**100.0%**	**104.7%**	**0.0%**	**0.0%**	**2.3%**	**0.06%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**0.1%**	**0.00%**	**0.0%**		**0.0%**	**0.00%**
Propensity to consume	**94.1%**	**93.2%**	**0.0%**	**0.0%**	**4.9%**	**0.04%**	**1.5%**	**-0.02%**	**1.5%**	**-0.02%**	**-0.6%**	**-0.01%**	**0.0%**		**0.3%**	
HH disposable income/GDP ratio	**92.3%**	**93.3%**	**0.0%**		**6.7%**		**1.1%**		**1.1%**		**2.0%**	**0.03%**	**0.0%**		**0.2%**	
PX/PM (Terms of trade)	**100.0%**	**98.9%**	**0.0%**		**3.8%**		**-0.3%**		**-0.3%**		**-0.6%**		**0.0%**		**0.0%**	

Table 5: Bayesian versus tentative estimates

	Tentative estimates 2006 YF (1)	% Difference tentative and conventional estimates	Bayesian estimates, 2006YF option 14 (2)	% Difference Bayesian and conventional estimates (3)	Prior Coefficient of Variation (4)	Posterior Coefficient of Variation option 14 (5)
Nr. of distorted values		*11*		*13*		*3*
GDP ECONOMIC ACTIVITIES, CONSTANT	**219,355,723,268**	**0.2%**	**218,918,539,581**	**0.0%**	**3.00%**	**0.07%**
GDP EXPENDITURES, CONSTANT	**217,212,744,301**	**-0.8%**	**218,918,539,581**	**0.0%**	**3.00%**	**0.07%**
Agriculture, forestry and fishing	25,909,930,997	0.2%	26,209,217,770	1.4%	12.00%	0.10%
Mining and quarrying	3,315,810,341	1.4%	3,265,850,268	-0.1%	12.00%	0.12%
Manufacturing	42,276,263,192	-1.9%	42,830,905,154	-0.6%	12.00%	0.26%
Electricity, gas and water supply	5,954,046,559	2.6%	5,942,627,996	2.4%	12.00%	0.36%
Construction	12,137,992,674	2.8%	10,890,498,320	-7.8%	12.00%	0.57%
Wholesale and retail trade	33,763,471,805	-0.3%	34,532,219,229	1.9%	12.00%	0.22%
Land transport	4,088,573,858	6.3%	3,904,397,446	1.5%	12.00%	0.47%
Transport (other), storage, telecommunication	11,344,152,665	0.1%	11,210,364,867	-1.1%	12.00%	0.29%
Financial and insurance activities	6,529,536,200	0.4%	6,482,752,094	-0.4%	1.00%	0.11%
Private services	58,130,636,364	0.1%	57,985,829,092	-0.1%	12.00%	0.02%
Public administration and defence; compulsory social security	14,664,192,249	0.0%	14,603,034,376	-0.4%	0.10%	0.04%
VALUE ADDED GROSS, TOTAL, CURRENT	**218,114,606,905**	**0.0%**	**217,857,696,610**	**-0.1%**	**12.00%**	**0.00%**
GDP by ECONOMIC ACTIVITIES, CURRENT	**229,859,500,562**	**0.0%**	**229,574,850,388**	**-0.1%**	**3.00%**	**0.01%**
Household final consumption	199,739,842,843	0.0%	205,142,806,802	2.7%	3.00%	0.00%
Government final consumption	19,237,195,482	0.0%	19,193,854,559	-0.2%	0.10%	0.01%
NPISH final consumption	1,939,433,975	-1.3%	1,976,245,004	0.5%	6.00%	0.45%
Gross fixed capital formation	46,019,465,063	-0.4%	43,740,744,068	-5.4%	3.00%	0.20%
Exports	57,302,114,005	0.0%	56,874,753,165	-0.7%	0.10%	0.09%
Minus: Imports	96,270,197,607	0.0%	97,353,553,210	1.1%	0.10%	0.05%
GDP EXPENDITURES, CURRENT	**228,140,677,515**	**-0.7%**	**229,574,850,388**	**-0.1%**	**3.00%**	**0.01%**
Compensation of employees	72,756,461,528	0.2%	73,112,111,229	0.7%	12.00%	0.14%
Other taxes less subsidies on production	1,372,143,961	9.9%	1,257,100,901	0.6%	0.10%	0.10%
Non-HH operating surplus, gross	76,980,086,024	0.6%	76,089,876,207	-0.6%	12.00%	0.18%
HH Operating surplus (on dwelling services), gross	19,504,504,390	0.1%	19,514,812,150	0.2%	12.00%	0.14%
Mixed Income, gross	47,580,064,856	-1.4%	47,883,796,123	-0.8%	12.00%	0.22%
VALUE ADDED GROSS, TOTAL, CURRENT	**218,114,606,905**	**0.0%**	**217,857,696,610**	**-0.1%**	**12.00%**	**0.00%**
HOUSEHOLD DISPOSABLE INCOME, GROSS	**221,122,537,153**	**3.1%**	**216,780,833,217**	**1.1%**	**12.00%**	**0.08%**
GDP growth rate 2006/2005	**5.597%**	**0.2%**	**5.4%**	**0.0%**		
Implicit price index Value Added, total	**104.80%**	**0.1%**	**104.7%**	**0.0%**	**0.10%**	**0.01%**
Propensity to consume	**94.08%**	**0.9%**	**94.6%**	**1.5%**	**6.00%**	**-0.02%**
HH disposable income/GDP ratio	**96.20%**	**2.9%**	**94.4%**	**1.1%**		

Table 6: Bayesian estimates based on complete versus partial data		Bayesian estimates in 2005YF with full vs. partial framework scope				Bayesian estimates in 2006YF with full vs. partial framework scope			
		Bayesian estimates of 2005GV option 1		Bayesian estimates of 2005YF option 12		Bayesian estimates of 2006YF option 14		Bayesian estimates of 2006GV option 13	
		% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nr. of distorted values		*0*	*3*	*15*	*3*	*13*	*3*	*5*	*5*
GDP ECONOMIC ACTIVITIES, CONSTANT		**0.0%**	**0.07%**	**0.3%**	**0.07%**	**0.0%**	**0.07%**	**-0.6%**	**0.07%**
GDP EXPENDITURES, CONSTANT		**0.0%**	**0.07%**	**0.3%**	**0.07%**	**0.0%**	**0.07%**	**-0.6%**	**0.07%**
Agriculture, forestry and fishing		0.0%	0.10%	0.1%	0.10%	1.4%	0.10%	-0.1%	0.10%
Mining and quarrying		0.0%	0.15%	0.0%	0.15%	-0.1%	0.12%	0.0%	0.12%
Manufacturing		0.0%	0.25%	1.2%	0.26%	-0.6%	0.26%	-1.0%	0.25%
Electricity, gas and water supply		0.0%	0.35%	0.3%	0.35%	2.4%	0.36%	-0.3%	0.36%
Construction		0.0%	0.53%	1.4%	0.54%	-7.8%	0.57%	-1.2%	0.52%
Wholesale and retail trade		0.0%	0.22%	-0.7%	0.22%	1.9%	0.22%	-0.7%	0.23%
Land transport		0.0%	0.46%	0.3%	0.48%	1.5%	0.47%	-0.4%	0.47%
Transport (other), storage, telecommunication		0.0%	0.28%	0.2%	0.29%	-1.1%	0.29%	-0.4%	0.30%
Financial and insurance activities		0.0%	0.03%	0.3%	0.12%	-0.4%	0.11%	0.0%	0.03%
Private services		0.0%	0.01%	0.3%	0.02%	-0.1%	0.02%	-0.2%	0.01%
Public administration and defence; compulsory social security		0.0%	0.02%	-0.5%	0.04%	-0.4%	0.04%	0.0%	0.02%
VALUE ADDED GROSS, TOTAL, CURRENT		**0.0%**	**0.00%**	**0.3%**	**0.00%**	**-0.1%**	**0.00%**	**-0.5%**	**0.00%**
GDP ECONOMIC ACTIVITIES, CURRENT		**0.0%**	**0.01%**	**0.3%**	**0.01%**	**-0.1%**	**0.01%**	**-0.4%**	**0.01%**
Household final consumption		0.0%	0.00%	-2.0%	0.00%	2.7%	0.00%	0.7%	0.00%
Government final consumption		0.0%	0.01%	-0.4%	0.01%	-0.2%	0.01%	0.0%	0.01%
NPISH final consumption		0.0%	0.42%	1.2%	0.45%	0.5%	0.45%	-0.2%	0.42%
Gross fixed capital formation		0.0%	0.20%	-8.5%	0.22%	-5.4%	0.20%	-4.0%	0.20%
Exports		0.0%	0.02%	5.4%	0.08%	-0.7%	0.09%	-0.1%	0.01%
Minus: Imports		0.0%	0.02%	-9.0%	0.05%	1.1%	0.05%	-0.2%	0.00%
GDP EXPENDITURES, CURRENT		**0.0%**	**0.01%**	**0.3%**	**0.01%**	**-0.1%**	**0.01%**	**-0.4%**	**0.01%**
Compensation of employees		0.0%	0.13%	0.4%	0.14%	0.7%	0.14%	-0.5%	0.14%
Other taxes less subsidies on production		0.0%	0.07%	0.0%	0.10%	0.6%	0.10%	0.0%	0.07%
Non-HH operating surplus, gross		0.0%	0.17%	-1.0%	0.18%	-0.6%	0.18%	-0.7%	0.18%
HH Operating surplus (on dwelling services), gross		0.0%	0.14%	0.1%	0.14%	0.2%	0.14%	-0.1%	0.14%
Mixed Income, gross		0.0%	0.21%	2.2%	0.21%	-0.8%	0.22%	-0.3%	0.21%
VALUE ADDED GROSS, TOTAL, CURRENT		**0.0%**	**0.00%**	**0.3%**	**0.00%**	**-0.1%**	**0.00%**	**-0.5%**	**0.00%**
HOUSEHOLD DISPOSABLE INCOME, GROSS	**2006 Prior values**	**0.0%**	**0.08%**	**-2.7%**	**0.08%**	**1.1%**	**0.08%**	**-0.1%**	**0.08%**
GDP growth rate 2006/2005	**5.4%**					**0.0%**		**-0.6%**	
Implicit price index Value Added, total	**104.7%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**0.0%**	**0.00%**
Propensity to consume	**93.2%**	**0.0%**	**-0.01%**	**0.7%**	**-0.02%**	**1.5%**	**-0.02%**	**0.7%**	**-0.02%**
HH disposable income/GDP ratio	**93.3%**	**0.0%**		**-2.8%**		**1.1%**		**0.3%**	

Table 7: Dependence of Bayesian estimates on framework scope and data availability

	Prior coefficients of variation	Bayesian estimates in 2005GV with full vs. partial framework scope						Bayesian estimates of alternatives, based on different data availability and precision, measured in 2006YF										
		Bayesian estimates of 2005GV option 1		SUT only option 9		IEA only option 10		Bayesian estimates of 2006YF High quality of data on -Household final consumption -Imports and exports -Financial corporations -Services option 20		Poor quality of household final consumption data and ratios option 16		Poor quality of export and import data option 17		Poor quality of data on services option 18		Pooor quality of financial corporations data option 19		Aggregate data option 21
		% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Nr. of distorted values		*0*	*3*	*0*	*6*	*2*	*15*	*11*	*2*	*13*	*2*	*15*	*9*	*32*	*2*	*23*	*4*	*34*
GDP ECONOMIC ACTIVITIES, CONSTANT	**3.00%**	**0.0%**	**0.07%**	**0.0%**	**0.10%**			**-1.1%**	**0.05%**	**-0.2%**	**0.07%**	**-0.3%**	**0.08%**	**2.3%**	**0.06%**	**-1.5%**	**0.06%**	**-1.6%**
GDP EXPENDITURES, CONSTANT	**3.00%**	**0.0%**	**0.07%**	**0.0%**	**0.10%**			**-1.1%**	**0.05%**	**-0.2%**	**0.07%**	**-0.3%**	**0.08%**	**2.3%**	**0.06%**	**-1.5%**	**0.06%**	**-1.6%**
Agriculture, forestry and fishing	12.00%	0.0%	0.10%	0.0%	0.41%			1.1%	0.10%	1.3%	0.10%	1.4%	0.10%	0.3%	0.10%	1.2%	0.10%	0.6%
Mining and quarrying	12.00%	0.0%	0.15%	0.0%	0.44%			-0.1%	0.12%	-0.1%	0.12%	-0.1%	0.12%	-0.3%	0.12%	-0.2%	0.12%	
Manufacturing	12.00%	0.0%	0.25%	0.0%	0.94%			-3.3%	0.23%	-1.0%	0.27%	-1.0%	0.27%	-10.9%	0.26%	-3.3%	0.24%	17.4%
Electricity, gas and water supply	12.00%	0.0%	0.35%	0.0%	0.67%			1.6%	0.36%	2.3%	0.36%	2.3%	0.36%	-0.2%	0.36%	1.5%	0.36%	
Construction	12.00%	0.0%	0.53%	0.0%	0.70%			-10.3%	0.57%	-8.1%	0.58%	-7.3%	0.57%	-25.1%	0.70%	-8.4%	0.56%	0.7%
Wholesale and retail trade	12.00%	0.0%	0.22%	0.0%	0.41%			0.5%	0.21%	1.7%	0.22%	0.9%	0.22%	67.5%	0.26%	-2.5%	0.22%	7.8%
Land transport	12.00%	0.0%	0.46%	0.0%	0.67%			0.6%	0.48%	1.1%	0.48%	1.2%	0.48%	4.0%	0.71%	0.7%	0.48%	
Transport (other), storage, telecommunication	12.00%	0.0%	0.28%	0.0%	0.66%			-2.0%	0.29%	-1.2%	0.29%	-1.1%	0.29%	-30.9%	0.88%	-1.7%	0.29%	
Financial and insurance activities	1.00%	0.0%	0.03%	0.0%	0.15%			-0.5%	0.11%	-0.4%	0.12%	-0.5%	0.12%	-2.7%	0.24%	-13.9%	0.75%	26.4%
Private services	12.00%	0.0%	0.01%	0.0%	0.36%			-0.6%	0.06%	-0.2%	0.09%	-0.2%	0.09%	-11.1%	0.15%	-0.5%	0.05%	
Public administration and defence; compulsory social security	0.10%	0.0%	0.02%	0.0%	0.06%			-1.4%	0.04%	-0.4%	0.04%	-1.4%	0.04%	-1.4%	0.04%	-1.5%	0.04%	0.0%
VALUE ADDED GROSS, TOTAL, CURRENT	**12.00%**	**0.0%**	**0.00%**	**0.0%**	**0.00%**			**-1.3%**	**0.00%**	**-0.3%**	**0.08%**	**-0.4%**	**0.08%**	**2.3%**	**0.00%**	**-2.0%**	**0.00%**	**-2.3%**
GDP ECONOMIC ACTIVITIES, CURRENT	**3.00%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**			**-1.3%**	**0.01%**	**-0.3%**	**0.08%**	**-0.4%**	**0.08%**	**2.2%**	**0.02%**	**-1.5%**	**0.01%**	**-2.1%**
Household final consumption	3.00%	0.0%	0.00%	0.0%	0.13%	0.1%	1.05%	1.0%	0.05%	2.6%	0.09%	0.1%	0.09%	3.1%	0.06%	0.6%	0.05%	1.3%
Government final consumption	0.10%	0.0%	0.01%	0.0%	0.04%	-2.4%	0.04%	1.2%	0.01%	-0.2%	0.03%	1.2%	0.03%	1.2%	0.01%	1.1%	0.01%	0.0%
NPISH final consumption	6.00%	0.0%	0.42%	0.0%	1.50%	-0.1%	1.79%	0.4%	0.44%	0.7%	0.47%	-0.4%	0.48%	1.3%	0.68%	-1.2%	0.45%	
Gross fixed capital formation	3.00%	0.0%	0.20%	0.0%	0.93%	0.1%	1.16%	-4.4%	0.19%	-5.6%	0.20%	-6.0%	0.21%	5.8%	0.19%	-5.5%	0.19%	-8.4%
Exports	0.10%	0.0%	0.02%	0.0%	0.08%	0.0%	0.09%	-0.8%	0.09%	-0.7%	0.09%	-8.5%	2.22%	-1.1%	0.09%	-0.5%	0.09%	0.0%
Minus: Imports	0.10%	0.0%	0.02%	0.0%	0.08%	0.0%	0.06%	1.1%	0.06%	1.1%	0.07%	-8.3%	1.31%	2.0%	0.06%	0.6%	0.06%	0.0%
GDP EXPENDITURES, CURRENT	**3.00%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**-0.1%**	**0.65%**	**-1.3%**	**0.01%**	**-0.3%**	**0.08%**	**-0.4%**	**0.08%**	**2.2%**	**0.02%**	**-1.5%**	**0.01%**	**-2.1%**
Compensation of employees	12.00%	0.0%	0.13%	0.0%	0.21%	0.2%	3.44%	-0.5%	0.13%	0.5%	0.14%	0.5%	0.14%	0.8%	0.13%	-1.0%	0.14%	-0.7%
Other taxes less subsidies on production	0.10%	0.0%	0.07%	0.0%	0.10%	-0.2%	0.02%	0.5%	0.10%	0.6%	0.10%	0.6%	0.10%	1.5%	0.09%	0.3%	0.10%	0.0%
Non-HH operating surplus, gross	12.00%	0.0%	0.17%	0.0%	0.17%	-0.2%	0.02%	-2.1%	0.17%	-0.8%	0.18%	-0.9%	0.18%	0.8%	0.17%	-3.2%	0.18%	-23.8%
HH Operating surplus (on dwelling services), gross	12.00%	0.0%	0.14%	0.0%	0.14%	0.0%	11.25%	0.1%	0.14%	0.2%	0.16%	0.1%	0.16%	-0.1%	0.14%	0.1%	0.14%	
Mixed Income, gross	12.00%	0.0%	0.21%	0.0%	0.32%	0.1%	6.42%	-2.0%	0.21%	-0.9%	0.22%	-1.2%	0.22%	8.1%	0.20%	-2.7%	0.21%	37.7%
VALUE ADDED GROSS, TOTAL, CURRENT	**12.00%**	**0.0%**	**0.00%**	**0.0%**	**0.00%**	**0.0%**	**1.29%**	**-1.3%**	**0.00%**	**-0.3%**	**0.08%**	**-0.4%**	**0.08%**	**2.3%**	**0.00%**	**-2.0%**	**0.00%**	**-2.3%**
HOUSEHOLD DISPOSABLE INCOME, GROSS	**12.00%**	**0.0%**	**0.08%**	**0.0%**	**0.00%**	**0.1%**	**1.01%**	**-0.3%**	**0.05%**	**1.0%**	**0.08%**	**-1.3%**	**0.09%**	**2.5%**	**0.06%**	**-0.7%**	**0.06%**	**4.2%**
GDP growth rate 2006/2005								**-1.2%**		**-0.2%**		**-0.4%**		**2.5%**		**-1.6%**		**-1.7%**
Implicit price index Value Added, total	**0.10%**	**0.0%**	**0.01%**	**0.0%**	**0.05%**			**-0.1%**	**0.01%**	**-0.1%**	**0.01%**	**-0.1%**	**0.01%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**0.0%**
Propensity to consume	**6.00%**	**0.0%**	**-0.01%**			**0.0%**	**0.09%**	**0.3%**	**0.03%**	**0.6%**	**0.03%**	**0.4%**	**0.04%**	**-0.3%**	**0.03%**	**0.4%**	**0.04%**	**-2.5%**
HH disposable income/GDP ratio		**0.0%**						**1.0%**		**1.2%**		**-0.8%**		**0.3%**		**0.8%**		**6.0%**

Table 8: Impact of prior reliability on Bayesian estimates and posterior reliability

	Prior coefficients of variation	Bayesian estimates in 2005GV: Standard prior coefficients of variation option 1		50% Increase of prior coefficients of variation option 3		50% Decrease of prior coefficients of variation option 2		Fictitious prior oefficients of variation (100%) option 4		Bayesian estimates in 2006YF: Standard prior coefficients of variation option 14		50% Increase of prior coefficients of variation option 6		Prior reliability of price indices poor (P) option 5	
		% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation	% Difference from conventional estimates	Posterior coefficient of variation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Nr. of distorted values		*0*	*3*	*0*	*11*	*0*	*4*	*0*	*64*	*13*	*3*	*13*	*5*	*10*	*11*
GDP ECONOMIC ACTIVITIES, CONSTANT	**3.00%**	**0.0%**	**0.07%**	**0.0%**	**0.16%**	**0.0%**	**0.05%**	**0.0%**	**8.18%**	**0.0%**	**0.07%**	**-0.4%**	**0.13%**	**0.8%**	**0.47%**
GDP EXPENDITURES, CONSTANT	**3.00%**	**0.0%**	**0.07%**	**0.0%**	**0.16%**	**0.0%**	**0.05%**	**0.0%**	**8.18%**	**0.0%**	**0.07%**	**-0.4%**	**0.13%**	**0.8%**	**0.47%**
Agriculture, forestry and fishing	12.00%	0.0%	0.10%	0.0%	0.42%	0.0%	0.14%	0.0%	10.80%	1.4%	0.10%	1.3%	0.19%	1.9%	0.12%
Mining and quarrying	12.00%	0.0%	0.15%	0.0%	0.46%	0.0%	0.15%	0.0%	5.55%	-0.1%	0.12%	-1.1%	0.23%	-1.7%	0.17%
Manufacturing	12.00%	0.0%	0.25%	0.0%	1.35%	0.0%	0.37%	0.0%	21.63%	-0.6%	0.26%	-1.0%	0.48%	-1.3%	0.30%
Electricity, gas and water supply	12.00%	0.0%	0.35%	0.0%	0.86%	0.0%	0.23%	0.1%	22.98%	2.4%	0.36%	2.7%	0.69%	2.6%	0.45%
Construction	12.00%	0.0%	0.53%	0.0%	1.08%	0.0%	0.28%	0.0%	30.42%	-7.8%	0.57%	-7.1%	1.10%	-9.7%	0.63%
Wholesale and retail trade	12.00%	0.0%	0.22%	0.0%	0.46%	0.0%	0.13%	0.0%	17.24%	1.9%	0.22%	0.1%	0.43%	-1.3%	0.27%
Land transport	12.00%	0.0%	0.46%	0.0%	0.96%	0.0%	0.25%	0.0%	28.80%	1.5%	0.47%	2.2%	0.90%	-0.8%	0.55%
Transport (other), storage, telecommunication	12.00%	0.0%	0.28%	0.0%	0.82%	0.0%	0.23%	0.0%	22.70%	-1.1%	0.29%	-1.2%	0.55%	-2.1%	0.39%
Financial and insurance activities	1.00%	0.0%	0.03%	0.0%	0.09%	0.0%	0.06%	0.0%	19.59%	-0.4%	0.11%	0.4%	0.18%	-1.3%	0.20%
Private services	12.00%	0.0%	0.01%	0.0%	0.43%	0.0%	0.12%	-0.1%	24.19%	-0.1%	0.02%	-0.2%	0.17%	-0.3%	0.03%
Public administration and defence; compulsory social security	0.10%	0.0%	0.02%	0.0%	0.04%	0.0%	0.04%	0.0%	9.35%	-0.4%	0.04%	-0.5%	0.04%	-0.5%	0.04%
VALUE ADDED GROSS, TOTAL, CURRENT	**12.00%**	**0.0%**	**0.00%**	**0.0%**	**0.14%**	**0.0%**	**0.00%**	**0.0%**	**4.02%**	**-0.1%**	**0.00%**	**-0.5%**	**0.14%**	**-1.0%**	**0.00%**
GDP ECONOMIC ACTIVITIES, CURRENT	**3.00%**	**0.0%**	**0.01%**	**0.0%**	**0.13%**	**0.0%**	**0.00%**	**0.0%**	**3.97%**	**-0.1%**	**0.01%**	**-0.4%**	**0.13%**	**-1.0%**	**0.01%**
Household final consumption	3.00%	0.0%	0.00%	0.0%	0.18%	0.0%	0.05%	0.0%	5.84%	2.7%	0.00%	2.2%	0.00%	1.9%	0.00%
Government final consumption	0.10%	0.0%	0.01%	0.0%	0.01%	0.0%	0.01%	0.0%	9.20%	-0.2%	0.01%	-0.3%	0.03%	-0.3%	0.01%
NPISH final consumption	6.00%	0.0%	0.42%	0.0%	0.81%	0.0%	0.22%	0.0%	6.92%	0.5%	0.45%	0.6%	0.87%	0.4%	0.46%
Gross fixed capital formation	3.00%	0.0%	0.20%	0.0%	1.76%	0.0%	0.45%	0.0%	24.81%	-5.4%	0.20%	-6.2%	0.39%	-5.9%	0.20%
Exports	0.10%	0.0%	0.02%	0.0%	0.05%	0.0%	0.05%	0.0%	15.25%	-0.7%	0.09%	-0.4%	0.09%	-0.7%	0.09%
Minus: Imports	0.10%	0.0%	0.02%	0.0%	0.04%	0.0%	0.03%	0.0%	10.07%	1.1%	0.05%	0.7%	0.06%	1.1%	0.05%
GDP EXPENDITURES, CURRENT	**3.00%**	**0.0%**	**0.01%**	**0.0%**	**0.13%**	**0.0%**	**0.00%**	**0.0%**	**3.97%**	**-0.1%**	**0.01%**	**-0.4%**	**0.13%**	**-1.0%**	**0.01%**
Compensation of employees	12.00%	0.0%	0.13%	0.0%	0.27%	0.0%	0.07%	0.0%	7.76%	0.7%	0.14%	0.4%	0.27%	-0.1%	0.15%
Other taxes less subsidies on production	0.10%	0.0%	0.07%	0.0%	0.07%	0.0%	0.06%	0.0%	4.30%	0.6%	0.10%	0.3%	0.10%	0.5%	0.10%
Non-HH operating surplus, gross	12.00%	0.0%	0.17%	0.0%	0.38%	0.0%	0.07%	0.0%	5.14%	-0.6%	0.18%	-1.2%	0.34%	-1.7%	0.19%
HH Operating surplus (on dwelling services), gross	12.00%	0.0%	0.14%	0.0%	0.29%	0.0%	0.06%	0.1%	28.61%	0.2%	0.14%	0.2%	0.30%	0.2%	0.13%
Mixed Income, gross	12.00%	0.0%	0.21%	0.0%	0.59%	0.0%	0.15%	0.0%	14.77%	-0.8%	0.22%	-0.8%	0.43%	-1.9%	0.22%
VALUE ADDED GROSS, TOTAL, CURRENT	**12.00%**	**0.0%**	**0.00%**	**0.0%**	**0.14%**	**0.0%**	**0.00%**	**0.0%**	**4.02%**	**-0.1%**	**0.00%**	**-0.5%**	**0.14%**	**-1.0%**	**0.00%**
HOUSEHOLD DISPOSABLE INCOME, GROSS	**12.00%**	**0.0%**	**0.08%**	**0.0%**	**0.16%**	**0.0%**	**0.05%**	**0.0%**	**5.59%**	**1.1%**	**0.08%**	**0.7%**	**0.15%**	**0.3%**	**0.09%**
GDP growth rate 2006/2005										**0.0%**		**-0.4%**		**0.8%**	
Implicit price index Value Added, total	**0.10%**	**0.0%**	**0.01%**	**0.0%**	**0.13%**	**0.0%**	**0.03%**	**0.0%**	**8.67%**	**0.0%**	**0.01%**	**0.0%**	**0.01%**	**-1.0%**	**0.17%**
Propensity to consume	**6.00%**	**0.0%**	**-0.01%**	**0.0%**	**0.07%**	**0.0%**	**0.02%**	**0.0%**	**1.75%**	**1.5%**	**-0.02%**	**1.4%**	**-0.03%**	**1.4%**	**-0.02%**
HH disposable income/GDP ratio		**0.0%**		**0.0%**		**0.0%**		**0.0%**		**1.1%**		**1.1%**		**1.2%**	
PX/PM (Terms of trade)		**0.0%**		**0.0%**		**0.0%**		**0.0%**		**-0.3%**		**-0.3%**		**-1.6%**	

Figure 1: BSCN framework, synoptic presentation

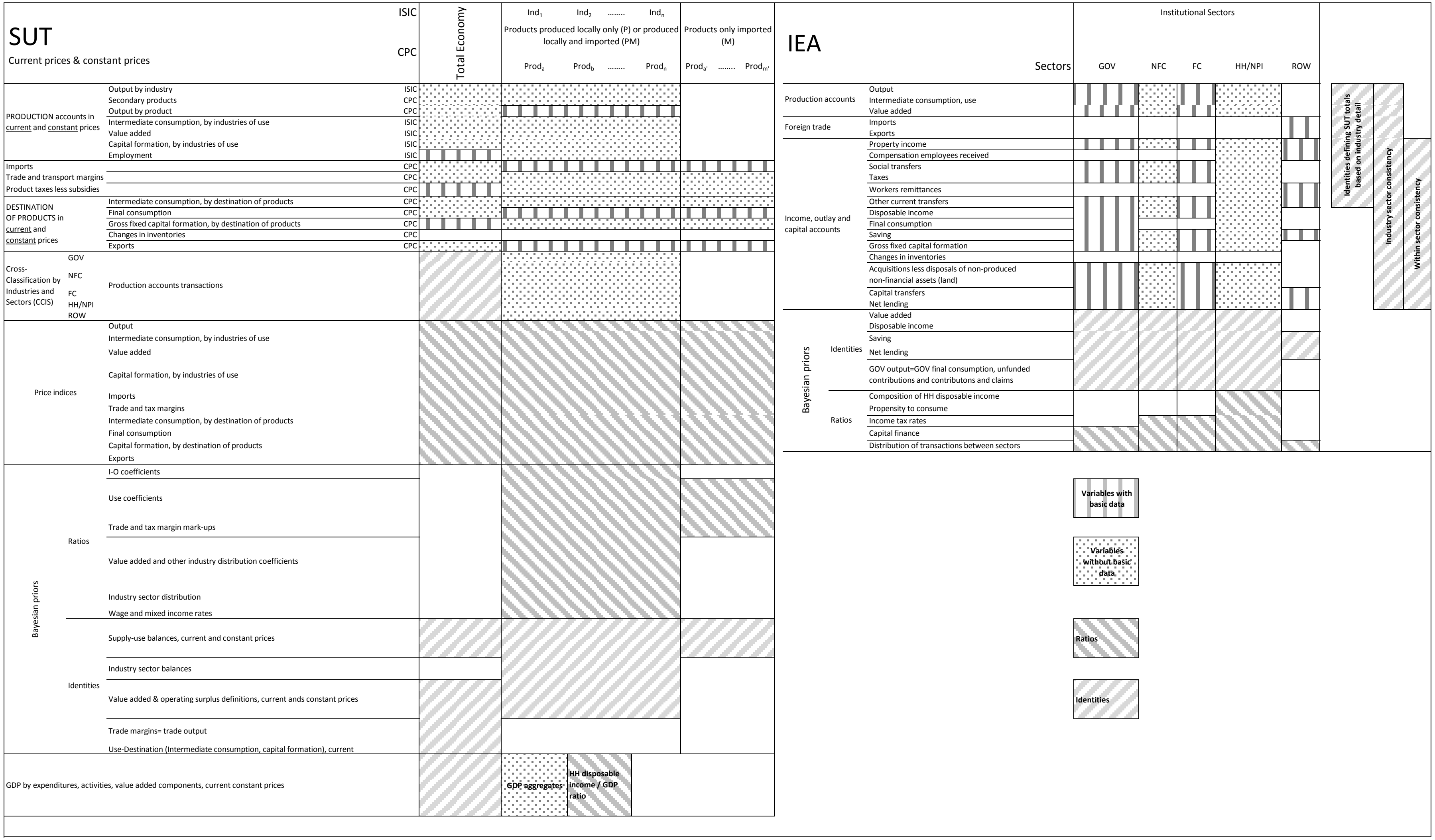
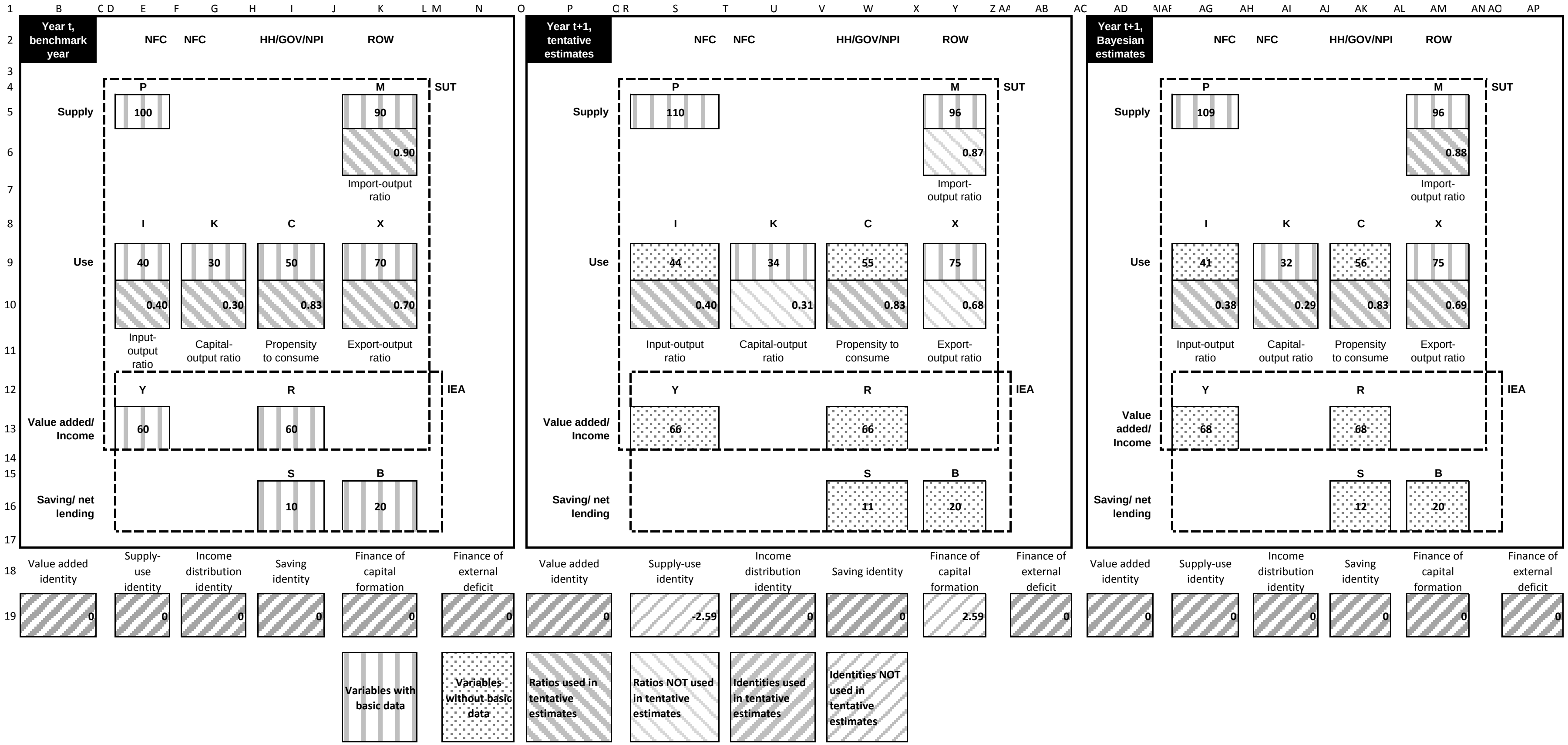


Figure 2: Comparison of tentative and Bayesian estimates in the BSNA framework, an example



Summary

This thesis deals with two issues: the design of frameworks of estimates for use in policy analysis and other analytical studies, and the development of a compilation methodology to arrive at estimates of variables within a framework. Such estimates should be comprehensive, covering all variables of the framework and not only those for which basic data are available, and also satisfying predefined consistency criteria. The resulting set of estimates can be used more effectively in policy analysis and analytical studies than if only limited inconsistent data sets were available.

The methodology used in the design of the frameworks and their compilation takes as its starting point the theories and practices of national accounting, but extends those to many other fields of study. The research reported in this thesis results in an operational instrument of framework design and compilation methodology that permits developing in a consistent manner comprehensive inter-disciplinary sets of estimates for many fields of study. The design of the frameworks and the variables included should be tailored to the specific analytical and policy issues addressed. An integral part of the operational instrument are two software programs that were designed in parallel with the methodology, i.e. the softwares of SNAER (System of National Accounts Estimation and Reconciliation) that generates the estimates, and of INTERFACE that carries out checks and then converts data in frameworks into formats that can be read by SNAER.

The work benefited much from the author's involvement in the development of the 1993 SNA (System of National Accounts) and the preparation of UNSD (United Nations Statistical Department) handbooks supporting its implementation. The design of the frameworks takes into account three different approaches to the development of national accounts frameworks:

1. the US approach to national accounting reflected in recent publications by Jorgenson, in which a top-down approach is favored in which major aggregates are linked and then further detailed,

2. the Stone sector approach reflected in the 1968 and 1993 SNA, in which the accounts are developed bottom-up, i.e. defining major aggregates as aggregations of micro unit data, and

3. the so-called micro-macro approach promoted by Ruggles and Ruggles proposes distinct sets of micro data to be maintained as the data bases for national accounts development, so that alternative frameworks supporting different types of analysis can be generated on the basis of the same micro unit data.

Furthermore comments by Vanoli, a prominent international expert on national accounting, on various aspects of national accounting, were extensively referred to in the thesis.

Two specific frameworks are included in the thesis: a framework developed for use in Central Banks in Central America (Chapter 6), and an economic-environmental accounts framework (Chapter 3). The framework developed for use in Central Banks in Central America, particularly Guatemala, is discussed in Chapter 6 and emphasizes the link with monetary policy. This framework, while mainly based on economic variables as included in the 1993 SNA, puts much emphasis on the use of the Integrated Economic Accounts (IEA) of the SNA, since that part of the SNA framework can be easily extended to financial accounts and balance sheets supporting monetary and fiscal policies. In the sectors identified in the IEA are the Government, the Central Bank and other banks and insurance schemes, as well as the Balance of Payments transactions with the Rest of the World. These are the three groups of sector transactions through which Central Banks and the Government influence the functioning of the economy. At present two extensions of this type of framework are contemplated. One is the use of a similar framework in quarterly accounts, thus including GDP related data of the SUT (Supply and Use Table) on economic development with IEA (Integrated Economic Accounts) data supporting monetary and fiscal policies; the other is the use of the framework for projections and impact analysis, simulating different types of responses by Central Banks and Governments to adverse economic developments.

The second framework included in the thesis is an economic-environmental accounts framework (Chapter 3), which is an essential component in the international standardization of environmental satellite accounts. In this framework the asset scope of the SNA is extended to include not only produced assets as covered in the SNA, but also natural assets, such as mineral resources, forests, water resources, fish stocks and even air. The framework is based on integrated SUT and asset balances. The SUT includes additionally imputed cost of depletion of mineral resources, forests and fish stock, and imputed cost of restoration or maintenance to respond to the degradation of land, forests and water and air through residuals emitted as a consequence of economic activities. In order to serve an alternative environment-oriented analysis, the framework includes environmental cost-adjusted concepts of value added and Domestic Product (EDP) and a concept called capital accumulation that reflects the changes in the stocks of assets between periods. The framework concepts are fully defined in terms of monetary values (market values or imputed values), but as imputed monetary valuation of environmental cost is still controversial, the framework may include, as proposed in the 2003 SEEA, a so-called 'hybrid' SUT, in which monetary values of SNA value added and cost, are combined with environmental costs of depletion and degradation that are expressed in terms of physical units without monetary valuation.

The author has also been involved in other policy-oriented frameworks, such as health accounts frameworks, and frameworks for regional development. The Introduction shows how a framework could be designed linking economic development of a country, reflected in GDP-related variables, with poverty analyses, reflected in variables such as the number of individuals below the poverty line. Through these framework extensions, demographic factors of population and employment are incorporated. In the design of frameworks, classifications and correspondences between those classifications play an important role. In Chapter 2 it is suggested that international classifications and correspondences be adjusted to national data availability and analytical needs, and that this be done at the most detailed level of each of the classifications that are used.

The compilation of the frameworks follow the principles of national accounting. In these practices, a limited set of data jointly with assumed relations (ratios) and identity links between the variables, are used to make estimates of a much larger set of variables than for which basic data are available. In a typical national accounts compilation of a Supply and Use Table (SUT) and Integrated Economic Accounts (IEA), estimates are made of approximately 2500 variables, while annually or quarterly only 20% of the variables are supported by basic data.

The essence of the approach to compilation developed in the thesis is that variables to be estimated, basic data available, assumptions in the form of ratios between variables, identities and also reliabilities of basic data and values of ratios are made explicit. The compilation approaches described in Chapters 4–6 thus formalize the conventional compilation approaches. This makes it easier to computerize the compilation, resulting in much quicker estimates than when made with help of the conventional approaches that include more manual components. There is also another difference with the conventional approaches and alternative computerized approaches such as ERETES (a national accounts software developed in cooperation between INSEE in France and EUROSTAT). The latter are based on what is called a sequential compilation approach, i.e. estimates are made of one group of variables, which are then used as inputs in the next step of the compilation. In the approaches developed in the thesis the sequential approach is replaced by a simultaneous compilation of variables, so that restrictions only incorporated in the last steps of the conventional compilation can influence the estimates already from the beginning.

The thesis includes two versions of the formalized compilation approach. The first is a linear programming (LP) approach presented in Chapter 4, in which the reliability is reflected in fixed intervals for basic data and ratio values. Those prior intervals are then used together with identities between variables and basic data, to arrive at posterior intervals (maxima and minima) of all variables, which are feasible within the prior restrictions. These posterior estimates of maxima and minima of variables are included in a large number of solution vectors, of which a simple average is made in order to arrive at posterior point estimates of the variables. In an alternative approach, developed later and presented in Chapters 5 and 6, Bayesian estimates are made, in which the fixed intervals of basic data and ratio values are replaced by the assumption that those values are lying within the range of a normal distribution. The shape of the normal distribution is determined by prior intervals of percent deviations that are allowed for the posterior values of the estimates to deviate from the prior values of the basic data and ratio values. Those percentage intervals, however, are not used as fixed intervals, but as coefficients of variation that are converted to the parameters of prior standard deviations and variances that determine the shape of the normal distribution of each basic data and ratio value. By minimizing the sum of the least squares of the prior values of basic data and ratio values minus the average posterior values to be estimated, under conditions defined by ratios and identities, average posterior estimated values are derived for each variable and ratio. By using the variances as weights, variables and ratios with a lower variance are adjusted less, and those with a higher variance are adjusted more, when changing prior to posterior values. This estimation method is reflected in the SNAER software.

From studies of several data frameworks that were compiled in the course of the thesis work, conclusions are reached with regard (i) the impact of basic data on the final estimates, (ii) the impact on final estimates of limited instead of comprehensive data sets, (iii) the need to develop so-called 'tentative estimates' based on a selection of identities and ratios prior to Bayesian or LP estimates, and (iv) the relation between the results of a framework approach in projections to the future and conventional modeling.